

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 29, 2004, 06:16:23 ; Search time 21 Seconds
(without alignments)
1190.943 Million cell updates/sec

Title: US-09-840-243C-11
 Perfect score: 1341
 Sequence: 1 MELTQPAEDLIQTQTPASE.....VIENHILKLFQSNLVPADPE 260

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 2833366

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```
Database :      PIR_78:*
1:  pir1:*
2:  pir2:*
3:  pir3:*
4:  pir4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	219.5	16.4	426	2	AE2149		hypothetical prote
2	214	16.0	2039	2	T15347		ankyrin-related un
3	207.5	15.5	1031	2	T43458		hypothetical prote
4	202.5	15.1	3924	2	S37431		ankyrin 2, neurona
5	193	14.4	4377	2	A55575		ankyrin 3, long sp
6	189	14.1	1001	2	S30385		G9a protein - huma
7	188.5	14.1	1765	2	T42714		ankyrin 3, splice
8	188.5	14.1	1940	2	T42715		ankyrin 3, splice
9	188.5	14.1	1943	2	T42713		ankyrin 3, splice
10	188.5	14.1	1961	2	T42716		ankyrin 3, splice
11	184	13.7	1881	1	SJHUK		ankyrin 1, erythro
12	183	13.6	1848	2	S37771		ankyrin, erythrocy
13	183	13.6	1862	2	I49502		ankyrin - mouse
14	182	13.6	1856	2	B35049		ankyrin 1, erythro
15	182	13.6	1880	2	A35049		ankyrin 1, erythro
16	181.5	13.5	1549	2	T13940		ankyrin - fruit fl
17	176.5	13.2	2584	2	T24158		hypothetical prote
18	176.5	13.2	2606	2	T24157		hypothetical prote
19	175	13.0	495	2	T27995		hypothetical prote
20	173.5	12.9	888	2	D84650		probable potassium
21	173	12.9	333	2	JC7713		ankyrin-repeat pro
22	172.5	12.9	1188	2	T19552		hypothetical prote
23	171.5	12.8	741	2	A45771		2-5A-dependent RNA
24	171	12.8	1398	2	T21884		hypothetical prote
25	169.5	12.6	1062	2	T14151		inv protein - mous
26	167	12.5	211	2	T18174		ankyrin repeat pro
27	166.5	12.4	1062	2	T30255		inversin - mouse
28	166	12.4	658	2	S68418		protein phosphatas
29	165.5	12.3	451	1	S68455		serine/threonine-s

ALIGNMENTS

RESULT 1

AE2149
hypothetical protein all2748 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AE2149
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata,
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ar
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AE2149
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-426 <KUR>
A;Cross-references: GB:BA000019; PIDN:BAB74447.1; PID:g17131841; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all2748

Query Match

Query Match	16.4%;	Score 219.5;	DB 2;	Length 426;
Best Local Similarity	30.7%;	Pred. No. 4.4e-10;		
Matches 62;	Conservative 34;	Mismatches 71;	Indels 35;	Gaps 4;

Qy	62	QAGSSLKHXSTTL-----TNRQRGNEVSALPATLDSLSIHQLAAQGELDQLK	107
Dd	234	QDGESALHLATVEGYVDVVQVLLNQGANQOIKNKLGDTPLLVAAQLQH-----DQIV	285
Qy	108	EHLRK-----GDNLVNKPDPERGFTPLIWASAFGEIETVRFLLEWGDADPHILAKERESAL	161
Dd	286	ETLLKYGANVHGDNL-----GETPLTLAASQGHATVRIILLDYGANANIPASDGKTAL	338
Qy	162	SLASTGGYTDIVGLLLERDVIDINIYDWNNGTPLLYAVRGNHVKVCVEALLARGADLTTEAD	221
Dd	339	IKATERNHPGVIQLLLAKGANVNYQDSVGATALIWAASGGYNKVQVILLEGADTNLKNR	398
Qy	222	SGYTPMDLAVALGYRKVQVQVIE	243
Dd	399	GGYTALMIAEENFGFRSIVQIILK	420

RESULT 2

ankyrin-related unc-44 - *Caenorhabditis elegans*
 T15347
 C/Species: *Caenorhabditis elegans*
 C/Date: 20-Sep-1999 #sequence revision 20-Sep-1999 #text_change 02-Aug-2002
 C/Accession: T15347; T15346; T15345; A57282; B57282; C57282
 R/Gattung, S.
 submitted to the EMBL Data Library, February 1996
 A/Description: The sequence of *C. elegans* cosmid B0350.
 A/Reference number: Z18332

Db 503 QQGASPNAAATSGYTPLHLAAREGHEDVAAP--LLDHGASLSITTKGFTPLHVAACYGK 560
QY 103 LDQLKEHLKRGDNLVNKPD-----ERGFT 126
Db 561 LEVASLLQKSAS---PDAAGKSGLTPLHVAHYDNQKVALLLLDQGASPHAAKNGYT 616
QY 127 PLIWASAFGEIETVRFLLEWGADPHILAKERESALSASTGGYTDIVGLLLERDVVDINIY 186
Db 617 PLHIAAKKNQMDIATSLLEYGADANAVTRQGIASVHLAAQEGHVDMSVLLSRNANVNLS 676
QY 187 DWNGGTPLLYAVRGNHVKCVALLARGADLTTEADSGYTPMDLAVALGYRKVQVQVIENHI 246
Db 677 NKSGLTPLHLAAQEDRVNVAEVLVNOGAHVDAQTKMGYTPLHVGYCHYGNIKIVNFFLQHS 736
QY 247 LKL 249
Db 737 AKV 739

RESULT 10
T42716
ankyryn 3, splice form 4 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C;Accession: T42716
R;Peters, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Yialamas, M.; Turtzo, L.
J. Cell Biol. 130, 313-330, 1995
A;Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene
the repeat domain.
A;Reference number: Z22237; MUID:95340633; PMID:7615634
A;Accession: T42716
A;Status: preliminary; translated from GE/EMBL/DBRJ
A;Molecule type: mRNA
A;Residues: 1-1961 <PET>
A;Cross-references: EMBL:L40632; NID:g710548; PID:g710552; PIDN:AAB01607.1
A;Experimental source: strain C57BL/6J; kidney
C;Genetics:
A;Gene: Ank3
A;Map position: 10
A;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing

Query Match 14.1%; Score 188.5; DB 2; Length 1961;
Best Local Similarity 25.5%; Pred. No. 1.3e-06;
Matches 62; Conservative 43; Mismatches 77; Indels 61; Gaps 6;
QY 62 QAGSSLLKHSTT-----LTNRQRGNEVSALPATLD---SLSI-----HQLAAQGE 102
Db 503 QQGASPNAAATSGYTPLHLAAREGHEDVAAP--LLDHGASLSITTKGFTPLHVAACYGK 560
QY 103 LDQLKEHLKRGDNLVNKPD-----ERGFT 126
Db 561 LEVASLLQKSAS---PDAAGKSGLTPLHVAHYDNQKVALLLLDQGASPHAAKNGYT 616
QY 127 PLIWASAFGEIETVRFLLEWGADPHILAKERESALSASTGGYTDIVGLLLERDVVDINIY 186
Db 617 PLHIAAKKNQMDIATSLLEYGADANAVTRQGIASVHLAAQEGHVDMSVLLSRNANVNLS 676
QY 187 DWNGGTPLLYAVRGNHVKCVALLARGADLTTEADSGYTPMDLAVALGYRKVQVQVIENHI 246
Db 677 NKSGLTPLHLAAQEDRVNVAEVLVNOGAHVDAQTKMGYTPLHVGYCHYGNIKIVNFFLQHS 736
QY 247 LKL 249
Db 737 AKV 739

RESULT 11
SJHUK
ankyryn 1, erythrocyte splice form 1 - human
N;Alternate names: ankyrin 2.1, erythrocyte; ankyrin-R
N;Contains: ankyrin 2.2
C;Species: Homo sapiens (man)

C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 22-Jun-1999
C;Accession: S08275; A33219; PC2220; A35443
R;Lux, S.E.; John, K.M.; Bennett, V.
Nature 344, 36-42, 1990
A;Title: Analysis of cDNA for human erythrocyte ankyrin indicates a repeated structure
A;Reference number: S08275; MUID:90158830; PMID:2137557
A;Accession: S08275
A;Molecule type: mRNA
A;Residues: 1-1881 <LUL>
A;Cross-references: EMBL:X16609; NID:g28701; PIDN:CAA34610.1; PID:g28702
A;Accession: A33219
A;Molecule type: protein
A;Residues: 2-7,'X',9-17,'X',19-20,'T',22-30;733-749,'A',751-753;828-833,'X',835-855,'Y'
X',1367;1383-1427;1601-1630;1686-1698,'D',1700;1763-1772 <LUX>
A;Note: 845-Arg and 1392-Thr were also found
R;Hermann, J.; Barel, M.; Frade, R.
Biochem. Biophys. Res. Commun. 204, 453-460, 1994
A;Title: Human erythrocyte ankyrin, a cytoskeleton component, generates the p57 membrai
A;Reference number: PC2220; MUID:95071348; PMID:7526850
A;Accession: PC2220
A;Molecule type: protein
A;Residues: 910-929 <HER>
R;Davis, L.H.; Bennett, V.
J. Biol. Chem. 265, 10589-10596, 1990
A;Title: Mapping the binding sites of human erythrocyte ankyrin for the anion exchanger
A;Reference number: A35443; MUID:90285190; PMID:2141335
A;Accession: A35443
A;Molecule type: protein
A;Residues: 'X',5,'X',7-12;403-417,'X',419-422,'H',424,'LQ';797-800,'L',802-814;862-863
C;Genetics:
A;Gene: GDB:ANK1; ANK
A;Cross-references: GDB:118737; OMIM:182900
A;Map position: 8p11.2-8p11.2
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing; phosphoprotein
F;2-1881/Product: ankyrin 1, erythrocyte form 1 #status predicted <MAT1>
F;2-1512,1675-1881/Product: ankyrin 2.2, erythrocyte #status predicted <MAT2>
F;2-827/Domain: 89K #status predicted <DOM1>
F;2-827/Region: anion exchange protein binding
F;44-76/Domain: ankyrin repeat homology <AN01>
F;77-109/Domain: ankyrin repeat homology <AN02>
F;110-142/Domain: ankyrin repeat homology <AN03>
F;143-171/Domain: ankyrin repeat homology <AN04>
F;172-204/Domain: ankyrin repeat homology <AN05>
F;205-237/Domain: ankyrin repeat homology <AN06>
F;238-270/Domain: ankyrin repeat homology <AN07>
F;271-303/Domain: ankyrin repeat homology <AN08>
F;304-336/Domain: ankyrin repeat homology <AN09>
F;337-369/Domain: ankyrin repeat homology <AN10>
F;370-402/Domain: ankyrin repeat homology <AN11>
F;403-435/Domain: ankyrin repeat homology <AN12>
F;436-468/Domain: ankyrin repeat homology <AN13>
F;469-501/Domain: ankyrin repeat homology <AN14>
F;502-534/Domain: ankyrin repeat homology <AN15>
F;535-567/Domain: ankyrin repeat homology <AN16>
F;568-600/Domain: ankyrin repeat homology <AN17>
F;601-633/Domain: ankyrin repeat homology <AN18>
F;634-666/Domain: ankyrin repeat homology <AN19>
F;667-699/Domain: ankyrin repeat homology <AN20>
F;700-732/Domain: ankyrin repeat homology <AN21>
F;733-765/Domain: ankyrin repeat homology <AN22>
F;766-798/Domain: ankyrin repeat homology <AN23>
F;828-1382/Domain: 62K #status predicted <DOM2>
F;828-1382/Region: spectrin binding
F;1383-1881/Domain: 55K #status predicted <DOM3>

Query Match 13.7%; Score 184; DB 1; Length 1881;
Best Local Similarity 23.1%; Pred. No. 2.8e-06;
Matches 54; Conservative 37; Mismatches 69; Indels 74; Gaps 3;
QY 74 TNRQRGNEVSALPATLDSLSIHQLAAQGEGLDQLKEHLKRGDNLVNKPDERGFTPLIWASA 133
Db 73 TTTTKKGN-----TALHTAALAGQDEVVELVNYGAN-VNAQSQKGFTPLVMAAQ 120

QY 134 FGEIETVRFLEWGADPHILAKERESALSASTGGYTDIVG----- 174
Db 121 ENHLEVVKFLLENGANQNVATEDGFTPLAVALQQGHENVVAHLINYGTKGKVRPALHIA 180
QY 175 -----LLLRDVEDINIYDNGG 191
Db 181 ARNDTRTAALLQNDPNPDVLSKTGFTPLHIAAHYENLNVAQLLNRGASVNFPTPQNGI 240
QY 192 TPLLYAVRGNHVKCEALLARGADLTTEADSGYTPMDLAVAGYRKVQVNIENH 245
Db 241 TPLHIASRRGNVIMVRLLLDRGAQIETRTKDELTPHCAARNGHVRISEILLDH 294

RESULT 12
S37771
Ankyrin, erythrocyte - mouse
C;Species: Mus musculus (house mouse)
C;Date: 27-May-1994 #sequence_revision 03-Aug-1995 #text_change 13-Aug-1999
C;Accession: S37771
R;Birckenmeier, C.S.; White, R.A.; Peters, L.L.; Hall, E.J.; Lux, S.E.; Barker, J.E.
J. Biol. Chem. 268, 9533-9540, 1993
A;Title: Complex patterns of sequence variation and multiple 5' and 3' ends are found an
A;Reference number: S37771; MUID:93252825; PMID:8486643
A;Accession: S37771
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1848 <BIR>
A;Cross-references: EMBL:X69063; NID:g311816; PIDN:CAA48801.1; PID:g311817
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing
F;48-80/Domain: ankyrin repeat homology <AN01>
F;81-113/Domain: ankyrin repeat homology <AN02>
F;114-146/Domain: ankyrin repeat homology <AN03>
F;147-175/Domain: ankyrin repeat homology <AN04>
F;176-208/Domain: ankyrin repeat homology <AN05>
F;209-241/Domain: ankyrin repeat homology <AN06>
F;242-274/Domain: ankyrin repeat homology <AN07>
F;275-307/Domain: ankyrin repeat homology <AN08>
F;308-340/Domain: ankyrin repeat homology <AN09>
F;341-373/Domain: ankyrin repeat homology <AN10>
F;374-406/Domain: ankyrin repeat homology <AN11>
F;407-439/Domain: ankyrin repeat homology <AN12>
F;440-472/Domain: ankyrin repeat homology <AN13>
F;473-505/Domain: ankyrin repeat homology <AN14>
F;506-538/Domain: ankyrin repeat homology <AN15>
F;539-571/Domain: ankyrin repeat homology <AN16>
F;572-604/Domain: ankyrin repeat homology <AN17>
F;605-637/Domain: ankyrin repeat homology <AN18>
F;638-670/Domain: ankyrin repeat homology <AN19>
F;671-703/Domain: ankyrin repeat homology <AN20>
F;704-736/Domain: ankyrin repeat homology <AN21>
F;737-769/Domain: ankyrin repeat homology <AN22>
F;770-802/Domain: ankyrin repeat homology <AN23>

Query Match 13.6%; Score 183; DB 2; Length 1848;
Best Local Similarity 23.1%; Pred. No. 3.3e-06;
Matches 54; Conservative 36; Mismatches 70; Indels 74; Gaps 3;
QY 74 TNRQRGNEVSALPATLDSLSIHQAQAGELDLQKEHLRKGDNLVNKPDERGFTPLIWASA 133
Db 77 TTTKKGN-----TALHIAALAGQDEVVRELNVYGAN-VNAQSQKGFTPLYMAAQ 124
QY 134 FGEIETVRFLEWGADPHILAKERESALSASTGGYTDIVG----- 174
Db 125 ENHLEVVKFLLENGANQNVATEDGFTPLAVALQQGHENVVAHLINYGTKGKVRPALHIA 184
QY 175 -----LLLRDVEDINIYDNGG 191
Db 185 ARNDTRTAALLQNDPNPDVLSKTGFTPLHIAAHYENLNVAQLLNRGASVNFPTPQNGI 244
QY 192 TPLLYAVRGNHVKCEALLARGADLTTEADSGYTPMDLAVAGYRKVQVNIENH 245

Db 245 TPLHIASRRGNVIMVRLLLDRGAQIETRTKDELTPHCAARNGHVRISEILLDH 298
RESULT 13
I49502
Ankyrin - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 13-Aug-1999
C;Accession: I49502
R;White, R.A.; Birkenmeier, C.S.; Peters, L.L.; Barker, J.E.; Lux, S.E.
Mamm. Genome 3, 281-285, 1992
A;Title: Murine erythrocyte ankyrin cDNA: Highly conserved regions of the regulatory do
A;Reference number: I49502; MUID:92345717; PMID:1386265
A;Accession: I49502
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1862 <RES>
A;Cross-references: GB:M84756; NID:g191939; PIDN:AAA37236.1; PID:g191940
C;Genetics:
A;Gene: Ank-1
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing
F;40-72/Domain: ankyrin repeat homology <AN01>
F;73-105/Domain: ankyrin repeat homology <AN02>
F;106-138/Domain: ankyrin repeat homology <AN03>
F;139-167/Domain: ankyrin repeat homology <AN04>
F;168-200/Domain: ankyrin repeat homology <AN05>
F;201-233/Domain: ankyrin repeat homology <AN06>
F;234-266/Domain: ankyrin repeat homology <AN07>
F;267-299/Domain: ankyrin repeat homology <AN08>
F;300-332/Domain: ankyrin repeat homology <AN09>
F;333-365/Domain: ankyrin repeat homology <AN10>
F;366-398/Domain: ankyrin repeat homology <AN11>
F;399-431/Domain: ankyrin repeat homology <AN12>
F;432-464/Domain: ankyrin repeat homology <AN13>
F;465-497/Domain: ankyrin repeat homology <AN14>
F;498-530/Domain: ankyrin repeat homology <AN15>
F;531-563/Domain: ankyrin repeat homology <AN16>
F;564-596/Domain: ankyrin repeat homology <AN17>
F;597-629/Domain: ankyrin repeat homology <AN18>
F;630-662/Domain: ankyrin repeat homology <AN19>
F;663-695/Domain: ankyrin repeat homology <AN20>
F;696-728/Domain: ankyrin repeat homology <AN21>
F;729-761/Domain: ankyrin repeat homology <AN22>
F;762-794/Domain: ankyrin repeat homology <AN23>

Query Match 13.6%; Score 183; DB 2; Length 1862;
Best Local Similarity 23.1%; Pred. No. 3.3e-06;
Matches 54; Conservative 36; Mismatches 70; Indels 74; Gaps 3;
QY 74 TNRQRGNEVSALPATLDSLSIHQAQAGELDLQKEHLRKGDNLVNKPDERGFTPLIWASA 133
Db 69 TTTKKGN-----TALHIAALAGQDEVVRELNVYGAN-VNAQSQKGFTPLYMAAQ 116
QY 134 FGEIETVRFLEWGADPHILAKERESALSASTGGYTDIVG----- 174
Db 117 ENHLEVVKFLLENGANQNVATEDGFTPLAVALQQGHENVVAHLINYGTKGKVRPALHIA 176
QY 175 -----LLLRDVEDINIYDNGG 191
Db 177 ARNDTRTAALLQNDPNPDVLSKTGFTPLHIAAHYENLNVAQLLNRGASVNFPTPQNGI 236
QY 192 TPLLYAVRGNHVKCEALLARGADLTTEADSGYTPMDLAVAGYRKVQVNIENH 245
Db 237 TPLHIASRRGNVIMVRLLLDRGAQIETRTKDELTPHCAARNGHVRISEILLDH 290

RESULT 14
B35049
Ankyrin 1, erythrocyte splice form 3 - human
N;Alternate names: ankyrin 2.1, erythrocyte; ankyrin-R
N;Contains: ankyrin 2.2, erythrocyte
C;Species: Homo sapiens (man)

C>Date: 17-Nov-1995 #sequence_revision 17-Nov-1995 #text_change 10-Jul-1998
C/Accession: B35049
R/Lambert, S.; Yu, H.; Prchal, J.T.; Lawler, J.; Ruff, P.; Speicher, D.; Cheung, M.C.;
Proc. Natl. Acad. Sci. U.S.A. 87, 1730-1734, 1990
A>Title: cDNA sequence for human erythrocyte ankyrin.
A/Reference number: A35049; MUID:90175370; PMID:1689849
A/Accession: B35049
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-1856 <LAM>
C/Genetics:
A/Gene: GDB:ANK1; ANK
A/Cross-references: GDB:118737; OMIM:182900
A/Map position: 8p11.2-8p11.2
C/Superfamily: ankyrin; ankyrin repeat homology
C/Keywords: alternative splicing
F;2-1856/Product: ankyrin 1, erythrocyte form 3 #status predicted <MAT>
F;2-1513,1676-1856/Product: ankyrin 2.2, erythrocyte #status predicted <MA2>
F;44-76/Domain: ankyrin repeat homology <AN01>
F;77-109/Domain: ankyrin repeat homology <AN02>
F;110-142/Domain: ankyrin repeat homology <AN03>
F;143-171/Domain: ankyrin repeat homology <AN04>
F;172-204/Domain: ankyrin repeat homology <AN05>
F;205-237/Domain: ankyrin repeat homology <AN06>
F;238-270/Domain: ankyrin repeat homology <AN07>
F;271-303/Domain: ankyrin repeat homology <AN08>
F;304-336/Domain: ankyrin repeat homology <AN09>
F;337-369/Domain: ankyrin repeat homology <AN10>
F;370-402/Domain: ankyrin repeat homology <AN11>
F;403-435/Domain: ankyrin repeat homology <AN12>
F;436-468/Domain: ankyrin repeat homology <AN13>
F;469-501/Domain: ankyrin repeat homology <AN14>
F;502-534/Domain: ankyrin repeat homology <AN15>
F;535-567/Domain: ankyrin repeat homology <AN16>
F;568-600/Domain: ankyrin repeat homology <AN17>
F;601-633/Domain: ankyrin repeat homology <AN18>
F;634-666/Domain: ankyrin repeat homology <AN19>
F;667-699/Domain: ankyrin repeat homology <AN20>
F;700-732/Domain: ankyrin repeat homology <AN21>
F;733-765/Domain: ankyrin repeat homology <AN22>
F;766-798/Domain: ankyrin repeat homology <AN23>

Query Match 13.6%; Score 182; DB 2; Length 1856;
Best Local Similarity 23.1%; Pred. No. 4e-06;
Matches 54; Conservative 37; Mismatches 69; Indels 74; Gaps 3;

QY 74 TNRQGNVSALPATLDSLSIHQLAAQGLDQLKEHLRGDNLVKNKPDGFTPLIWASA 133
Db 73 TTTKGN-----TALHIAALAGQDEVVRELVNYGAN-VNAQSQKGFTPLYMAAQ 120
QY 134 FGEIETVRFLEWGADPHILAKERESALSASTGGYTDIVG----- 174
Db 121 ENHLEVVKFELLENGANQNVATEDGFTPLAVALQQGHENVVAHLINYGTKGVRLPALHIA 180
QY 175 -----LLLRDVDINIYDNGG 191
Db 181 ARNDDTRTAAVLLQNDPNPDVLSKTGFTPLHIAAHYENLNVAQLLNRGSSVNFPTQNGI 240
QY 192 TPLLYAVRGNHVKCVEALLARGADLTTEADSGYTPMDLAVALGYRKVQVNIENH 245
Db 241 TPLHIASRRGNVIMVRLLLDRGAQIETKTDELTPHCAARNGHVRISILLDH 294

RESULT 15
A35049
N;Alternate names: ankyrin 2.1, erythrocyte; ankyrin-R
N;Contains: ankyrin 2.2, erythrocyte
C;Species: Homo sapiens (man)
C/Date: 27-Jul-1990 #sequence_revision 01-Oct-1992 #text_change 04-Sep-1998
C/Accession: A35049
R/Lambert, S.; Yu, H.; Prchal, J.T.; Lawler, J.; Ruff, P.; Speicher, D.; Cheung, M.C.;
Proc. Natl. Acad. Sci. U.S.A. 87, 1730-1734, 1990

A>Title: cDNA sequence for human erythrocyte ankyrin.
A/Reference number: A35049; MUID:90175370; PMID:1689849
A/Accession: A35049
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-1880 <LAM>
A/Cross-references: GB:M28880
C/Genetics:
A/Gene: GDB:ANK1; ANK
A/Cross-references: GDB:118737; OMIM:182900
A/Map position: 8p11.2-8p11.2
C/Superfamily: ankyrin; ankyrin repeat homology
C/Keywords: alternative splicing; cytoskeleton
F;2-1880/Product: ankyrin 1, erythrocyte form 2 #status predicted <MAT>
F;2-1513,1676-1880/Product: ankyrin 2.2, erythrocyte #status predicted <MA2>
F;44-76/Domain: ankyrin repeat homology <AN01>
F;77-109/Domain: ankyrin repeat homology <AN02>
F;110-142/Domain: ankyrin repeat homology <AN03>
F;143-171/Domain: ankyrin repeat homology <AN04>
F;172-204/Domain: ankyrin repeat homology <AN05>
F;205-237/Domain: ankyrin repeat homology <AN06>
F;238-270/Domain: ankyrin repeat homology <AN07>
F;271-303/Domain: ankyrin repeat homology <AN08>
F;304-336/Domain: ankyrin repeat homology <AN09>
F;337-369/Domain: ankyrin repeat homology <AN10>
F;370-402/Domain: ankyrin repeat homology <AN11>
F;403-435/Domain: ankyrin repeat homology <AN12>
F;436-468/Domain: ankyrin repeat homology <AN13>
F;469-501/Domain: ankyrin repeat homology <AN14>
F;502-534/Domain: ankyrin repeat homology <AN15>
F;535-567/Domain: ankyrin repeat homology <AN16>
F;568-600/Domain: ankyrin repeat homology <AN17>
F;601-633/Domain: ankyrin repeat homology <AN18>
F;634-666/Domain: ankyrin repeat homology <AN19>
F;667-699/Domain: ankyrin repeat homology <AN20>
F;700-732/Domain: ankyrin repeat homology <AN21>
F;733-765/Domain: ankyrin repeat homology <AN22>
F;766-798/Domain: ankyrin repeat homology <AN23>

Query Match 13.6%; Score 182; DB 2; Length 1880;
Best Local Similarity 23.1%; Pred. No. 4.1e-06;
Matches 54; Conservative 37; Mismatches 69; Indels 74; Gaps 3;

QY 74 TNRQGNVSALPATLDSLSIHQLAAQGLDQLKEHLRGDNLVKNKPDGFTPLIWASA 133
Db 73 TTTKGN-----TALHIAALAGQDEVVRELVNYGAN-VNAQSQKGFTPLYMAAQ 120
QY 134 FGEIETVRFLEWGADPHILAKERESALSASTGGYTDIVG----- 174
Db 121 ENHLEVVKFELLENGANQNVATEDGFTPLAVALQQGHENVVAHLINYGTKGVRLPALHIA 180
QY 175 -----LLLRDVDINIYDNGG 191
Db 181 ARNDDTRTAAVLLQNDPNPDVLSKTGFTPLHIAAHYENLNVAQLLNRGSSVNFPTQNGI 240
QY 192 TPLLYAVRGNHVKCVEALLARGADLTTEADSGYTPMDLAVALGYRKVQVNIENH 245
Db 241 TPLHIASRRGNVIMVRLLLDRGAQIETKTDELTPHCAARNGHVRISILLDH 294

Search completed: April 29, 2004, 06:27:18
Job time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 29, 2004, 06:24:08 ; Search time 23 Seconds
(without alignments)
583.598 Million cell updates/sec

Title: US-09-840-243C-11
Perfect score: 1341
Sequence: 1 MELTQPAEDLIQTQTTPASE.....VIENHILKLFQSNLVPADPE 260

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1341	100.0	260	2	US-09-172-977-1
2	1341	100.0	260	4	US-09-404-108-1
3	1113.5	83.0	269	4	US-09-339-746A-2
4	205	15.3	348	2	US-09-031-485-28
5	205	15.3	348	2	US-08-847-429A-28
6	205	15.3	348	3	US-09-065-474-28
7	205	15.3	348	4	US-09-557-034-28
8	205	15.3	1745	2	US-09-031-485-33
9	205	15.3	1745	2	US-08-847-429A-33
10	205	15.3	1745	3	US-09-065-474-33
11	205	15.3	1745	4	US-09-557-034-33
12	202.5	15.1	1839	2	US-09-172-977-4
13	202.5	15.1	1839	4	US-09-404-108-4
14	197	14.7	843	2	US-09-172-977-3
15	197	14.7	843	4	US-09-404-108-3
16	187.5	14.0	1088	3	US-09-082-059-2
17	187	13.9	673	3	US-09-196-387-8
18	187	13.9	673	4	US-09-841-835-8
19	187	13.9	949	3	US-09-196-387-10
20	187	13.9	949	4	US-08-841-835-10
21	187	13.9	1327	3	US-09-196-387-2
22	187	13.9	1327	4	US-09-841-835-2
23	187	13.9	1327	4	US-09-972-115A-8
24	179	13.3	994	4	US-10-164-595-38
25	175	13.0	33	4	US-09-339-746A-3
26	175	13.0	435	4	US-09-533-029-56
27	172.5	12.9	1619	4	US-09-392-812A-4

28	171.5	12.8	741	2	US-08-462-481-2	Sequence 2, Appli
29	171.5	12.8	741	2	US-08-436-771-2	Sequence 2, Appli
30	171.5	12.8	741	2	US-08-434-998-2	Sequence 2, Appli
31	171.5	12.8	741	2	US-08-487-797-2	Sequence 2, Appli
32	171.5	12.8	741	2	US-08-701-005A-2	Sequence 2, Appli
33	171.5	12.8	741	2	US-08-479-895-2	Sequence 2, Appli
34	171.5	12.8	741	3	US-08-943-956A-2	Sequence 2, Appli
35	171.5	12.8	741	5	PCT-US95-02058-2	Sequence 2, Appli
36	170.5	12.7	228	4	US-09-543-681A-4781	Sequence 4781, Ap
37	169.5	12.6	741	2	US-08-436-771-4	Sequence 4, Appli
38	169.5	12.6	741	2	US-08-434-998-4	Sequence 4, Appli
39	169.5	12.6	741	2	US-08-487-797-4	Sequence 4, Appli
40	169.5	12.6	741	5	PCT-US95-02058-4	Sequence 4, Appli
41	166.5	12.4	248	4	US-09-328-352-6520	Sequence 6520, Ap
42	165.5	12.3	345	4	US-09-252-991A-17809	Sequence 17809, A
43	165.5	12.3	452	3	US-09-035-706-2	Sequence 2, Appli
44	165.5	12.3	452	3	US-08-955-841-2	Sequence 2, Appli
45	165.5	12.3	452	4	US-09-390-425-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-172-977-1
; Sequence 1, Application US/09172977
; Patent No. 5989863
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Yue, Henry
; TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN
; FILE REFERENCE: PF-0615 US
; CURRENT APPLICATION NUMBER: US/09/172,977
; CURRENT FILING DATE: 1998-10-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 1808075
US-09-172-977-1

Query Match	100.0%;	Score	1341;	DB	2;	Length	260;
Best Local Similarity	100.0%;	Pred. No.	1.2e-129;				
Matches	260;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
QY	1	MELTQPAEDLIQTQT	PASELGDPDPGEEA	ADGSDTVVLSLFPCTPEPVNPEPDASVSS	60		
Db	1	MELTQPAEDLIQTQT	PASELGDPDPGEEA	ADGSDTVVLSLFPCTPEPVNPEPDASVSS	60		
QY	61	POAGSSLKHSHTTL	TNRQRNEVSALPAT	LDLSLSIHQLAAQGELDQLKEHLRKGDNLVNKP	120		
Db	61	POAGSSLKHSHTTL	TNRQRNEVSALPAT	LDLSLSIHQLAAQGELDQLKEHLRKGDNLVNKP	120		
QY	121	DERGFTPLIWASAF	GEIETVRFLLEWGADPHILAKERESALS	LASTGGYTDIVGLLLERD	180		
Db	121	DERGFTPLIWASAF	GEIETVRFLLEWGADPHILAKERESALS	LASTGGYTDIVGLLLERD	180		
QY	181	VDINIYDWNNGT	PLLAVRGNHVKC	VEALLARGADLTTEADSGYTPMDLAVALGYRKVQQ	240		
Db	181	VDINIYDWNNGT	PLLAVRGNHVKC	VEALLARGADLTTEADSGYTPMDLAVALGYRKVQQ	240		
QY	241	VIENHILKLFQSN	LVLPADPE	260			
Db	241	VIENHILKLFQSN	LVLPADPE	260			

RESULT 2
US-09-404-108-1

; Sequence 1, Application US/09404108
; Patent No. 6590077
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Yue, Henry
; TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN
; FILE REFERENCE: PF-0615 US
; CURRENT APPLICATION NUMBER: US/09/404,108
; CURRENT FILING DATE: 1999-09-23
; EARLIER APPLICATION NUMBER: 09/172,977
; EARLIER FILING DATE: 1998-10-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 1808075
US-09-404-108-1

Query Match 100.0%; Score 1341; DB 4; Length 260;
Best Local Similarity 100.0%; Pred. No. 1.2e-129;
Matches 260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MELTQPAEDLIQTQTPASELGDPEEADGSDTVVLSLFPCTPEPVNPEPDASVSS 60
Db 1 MELTQPAEDLIQTQTPASELGDPEEADGSDTVVLSLFPCTPEPVNPEPDASVSS 60
QY 61 PQAGSLKHSTTLNRQRGNEVSALPATLDSLSIHQLAAQGHLDQLKEHLRKGDNLVNKP 120
Db 61 PQAGSLKHSTTLNRQRGNEVSALPATLDSLSIHQLAAQGHLDQLKEHLRKGDNLVNKP 120
QY 121 DERGFTPLIWASAFGEIETVRFLEWGADPHILAKERESALSASTGGYTDIVGLLLEED 180
Db 121 DERGFTPLIWASAFGEIETVRFLEWGADPHILAKERESALSASTGGYTDIVGLLLEED 180
QY 181 VDINIYDWNGGTPLLYAVRGNHVKCVALLARGADLTTEADSGYTPMDLAVALGYRKVQQ 240
Db 181 VDINIYDWNGGTPLLYAVRGNHVKCVALLARGADLTTEADSGYTPMDLAVALGYRKVQQ 240
QY 241 VIENHILKLFQSNLVPADPE 260
Db 241 VIENHILKLFQSNLVPADPE 260

RESULT 3
US-09-339-746A-2
; Sequence 2, Application US/09339746A
; Patent No. 6617427
; GENERAL INFORMATION:
; APPLICANT: Tschlis, Philip N
; APPLICANT: Makris, Antonios
; TITLE OF INVENTION: Nucleic Acid Molecule Encoding An Ankyrin Repeat Tvl-1 Protein An
; FILE REFERENCE: 99-40123US
; CURRENT APPLICATION NUMBER: US/09/339,746A
; CURRENT FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: US 60/090,742
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Amino acid sequence of the murine Tvl-1 protein
US-09-339-746A-2
Query Match 83.0%; Score 1113.5; DB 4; Length 269;

Best Local Similarity 81.9%; Pred. No. 3.2e-106;
Matches 221; Conservative 15; Mismatches 23; Indels 11; Gaps 2;
QY 1 MELTQPAEDLIQTQTPASELGDPEEADGSDTVVLSLFPCTPEPVNPEPDASVSS 60
Db 1 MEPTQVAENLVNQPPVDPLEDPEDTRDESPENSDTVVLSLFPCTPDVNPEDASASS 60
QY 61 PQAGSLKHSTTLNRQRGNEVSALPATLDSLSIHQLAAQGHLDQLKEHLRK----- 112
Db 61 LQ-GSFLKHSTTLNRQRGNEVSALPATLDSLSIHQLAAQGHLSQLKHLRKGACPACTC 119
QY 113 --GDNLVNKPDERGFTPLIWASAFGEIETVRFLEWGADPHILAKERESALSASTGGYT 170
Db 120 LSGNNLINKPDERGFTPLIWASAFGEIETVRFLEWGADPHILAKERESALSASTGGYT 179
QY 171 DIVGLLLEEDVDINIYDWNGGTPLLYAVRGNHVKCVALLARGADLTTEADSGYTPMDLA 230
Db 180 DIVRLLLDRDVIDINIYDWNGGTPLLYAVRGNHVKCVALLARGADLTTEADSGYTPMDLA 239
QY 231 VALGYRKVQVVIENHILKLFQSNLVPADPE 260
Db 240 VALGYRKVQVVMESHILRLFQSTLGPVDPE 269

RESULT 4
US-09-031-485-28
; Sequence 28, Application US/09031485
; Patent No. 5824306
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,485
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,429
; FILING DATE: 24-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 348 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-031-485-28

Query Match 15.3%; Score 205; DB 2; Length 348;
Best Local Similarity 33.9%; Pred. No. 1.2e-12;
Matches 58; Conservative 28; Mismatches 77; Indels 8; Gaps 3;


```

;
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/557,034
; FILING DATE: 21-Apr-2000
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/065,474
; FILING DATE: <Unknown>
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
;
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 348 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 28:
US-09-557-034-28

```

```

Query Match 15.3%; Score 205; DB 4; Length 348;
Best Local Similarity 33.9%; Pred. No. 1.2e-12;
Matches 58; Conservative 28; Mismatches 77; Indels 8; Gaps 3;

QY 75 NRQRGNEVSALPATLDSLSIHQAAQGLDQLKEHLRKGDNLVNKPDERGFTPLIWASAF 134
Db 28 NSQHSNKGES-----SASFLLRAARAGNLDRLVLELLRSSTD-INTCNANGLNALHLASKE 80

QY 135 GEIETVRFLEWGADPHILAKERESALSASTGGYTDIVGLLLERDVIDINIYDWNNGTPTL 194
Db 81 GHHEVVRELLKRAKADVDAAATRKGNTHALHSLAGQELIVTVLVENGANVNVQSLNGFTPL 140

QY 195 LYAVRGNHVKCVALLARGADLTTEADSGYTPMDLAVALGY-RKVQQVNIEN 244
Db 141 YMAAQENHESVVRVYLLAHNANQALSTEDGFTPLAVALQQGHDRVAVLLEN 191

```

```

RESULT 8
US-09-031-485-33
; Sequence 33, Application US/09031485
; Patent No. 5824306
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,485
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,429
; FILING DATE: 24-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
;
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1745 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
; US-09-031-485-33

```

```

Query Match 15.3%; Score 205; DB 2; Length 1745;
Best Local Similarity 33.9%; Pred. No. 1.4e-11;
Matches 58; Conservative 28; Mismatches 77; Indels 8; Gaps 3;

QY 75 NRQRGNEVSALPATLDSLSIHQAAQGLDQLKEHLRKGDNLVNKPDERGFTPLIWASAF 134
Db 28 NSQHSNKGES-----SASFLLRAARAGNLDRLVLELLRSSTD-INTCNANGLNALHLASKE 80

QY 135 GEIETVRFLEWGADPHILAKERESALSASTGGYTDIVGLLLERDVIDINIYDWNNGTPTL 194
Db 81 GHHEVVRELLKRAKADVDAAATRKGNTHALHSLAGQELIVTVLVENGANVNVQSLNGFTPL 140

QY 195 LYAVRGNHVKCVALLARGADLTTEADSGYTPMDLAVALGY-RKVQQVNIEN 244
Db 141 YMAAQENHESVVRVYLLAHNANQALSTEDGFTPLAVALQQGHDRVAVLLEN 191

```

```

RESULT 9
US-08-847-429A-33
; Sequence 33, Application US/08847429A
; Patent No. 5827692
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,429A
; FILING DATE: 24-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459

```

REFERENCE/DOCKET NUMBER: HW-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 970/493-7272
TELEFAX: 970/484-9505
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 1745 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-847-429A-33

Query Match 15.3%; Score 205; DB 2; Length 1745;
Best Local Similarity 33.9%; Pred. No. 1.4e-11;
Matches 58; Conservative 28; Mismatches 77; Indels 8; Gaps 3;

QY 75 NRQRGNEVSALPATLDSLSIHQAAQGGELDQLKEHLRKGDNLVNKPDERGFTPLIWASAF 134
Db 28 NSQHSNKGES-----SASFRAARAGNLDRLVLELLRSSTD-INTCNANGNLNHLASKE 80
QY 135 GEIETVRFLLEWGADPHILAKERESALSASTGGYTDIVGLLLERDLDVLDINIYDWNNGTPL 194
Db 81 GHHEVVRELLKRAKADVDATRKGNLTALHIASLAGQELIVTVLVENGANVNVQSLNGFTPL 140
QY 195 LYAVRGNHVKCVALLARGADLTTEADSGYTPMDLAVAGY-RKVQQVNIEN 244
Db 141 YMAAQENHESVVRVYLLAHNANQALSTEDGFTPLAVALQQGHDRVAVLLEN 191

RESULT 10
US-09-065-474-33
Sequence 33, Application US/09065474
Patent No. 6063599
GENERAL INFORMATION:
APPLICANT: Tang, Liang
APPLICANT: Blehm, E. Scot
TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
TITLE OF INVENTION: USES THEREOF
NUMBER OF SEQUENCES: 171
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carol Talkington Verser, Ph.D.
ADDRESSEE: Heska Corporation
STREET: 1825 Sharp Point Drive
CITY: Fort Collins
STATE: Colorado
COUNTRY: USA
ZIP: 80525
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: WordPerfect for Windows, Version 7.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065,474
FILING DATE: 24-APR-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Verser, Carol Talkington
REGISTRATION NUMBER: 37,459
REFERENCE/DOCKET NUMBER: HW-5-C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 970/493-7272
TELEFAX: 970/484-9505
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 1745 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-065-474-33

Query Match 15.3%; Score 205; DB 3; Length 1745;

Best Local Similarity 33.9%; Pred. No. 1.4e-11;
Matches 58; Conservative 28; Mismatches 77; Indels 8; Gaps 3;
QY 75 NRQRGNEVSALPATLDSLSIHQAAQGGELDQLKEHLRKGDNLVNKPDERGFTPLIWASAF 134
Db 28 NSQHSNKGES-----SASFRAARAGNLDRLVLELLRSSTD-INTCNANGNLNHLASKE 80
QY 135 GEIETVRFLLEWGADPHILAKERESALSASTGGYTDIVGLLLERDLDVLDINIYDWNNGTPL 194
Db 81 GHHEVVRELLKRAKADVDATRKGNLTALHIASLAGQELIVTVLVENGANVNVQSLNGFTPL 140
QY 195 LYAVRGNHVKCVALLARGADLTTEADSGYTPMDLAVAGY-RKVQQVNIEN 244
Db 141 YMAAQENHESVVRVYLLAHNANQALSTEDGFTPLAVALQQGHDRVAVLLEN 191

RESULT 11
US-09-557-034-33
Sequence 33, Application US/09557034
Patent No. 6365569
GENERAL INFORMATION:
APPLICANT: Tang, Liang
APPLICANT: Blehm, E. Scot
TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
TITLE OF INVENTION: USES THEREOF
NUMBER OF SEQUENCES: 171
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carol Talkington Verser, Ph.D.
ADDRESSEE: Heska Corporation
STREET: 1825 Sharp Point Drive
CITY: Fort Collins
STATE: Colorado
COUNTRY: USA
ZIP: 80525
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: WordPerfect for Windows, Version 7.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/557,034
FILING DATE: 21-Apr-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/065,474
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Verser, Carol Talkington
REGISTRATION NUMBER: 37,459
REFERENCE/DOCKET NUMBER: HW-5-C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 970/493-7272
TELEFAX: 970/484-9505
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 1745 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 33:
US-09-557-034-33

Query Match 15.3%; Score 205; DB 4; Length 1745;
Best Local Similarity 33.9%; Pred. No. 1.4e-11;
Matches 58; Conservative 28; Mismatches 77; Indels 8; Gaps 3;
QY 75 NRQRGNEVSALPATLDSLSIHQAAQGGELDQLKEHLRKGDNLVNKPDERGFTPLIWASAF 134
Db 28 NSQHSNKGES-----SASFRAARAGNLDRLVLELLRSSTD-INTCNANGNLNHLASKE 80
QY 135 GEIETVRFLLEWGADPHILAKERESALSASTGGYTDIVGLLLERDLDVLDINIYDWNNGTPL 194

Db 81 GHHEVVRELLKRXKADVDAAATRKGNHTALHIASLAGQELIVTVLVNGANVNVQSLNGFTPL 140

QY 195 LYAVRGNHVKVEALLARGADLTTEADSGYTPMDLAVAGY-RKVQQVNIEN 244

Db 141 YMAAQENHESVRYLLAHNANQALSTEDGFTPLAVALQQGHDRVVAVLLEN 191

RESULT 12

US-09-172-977-4

Query Match 15.1%; Score 202.5; DB 2; Length 1839;

Best Local Similarity 29.3%; Pred. No. 2.8e-11;

Matches 56; Conservative 43; Mismatches 81; Indels 11; Gaps 4;

GENERAL INFORMATION: US/09172977

Patent No. 5989863

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Guegler, Karl J.

APPLICANT: Corley, Neil C.

APPLICANT: Yue, Henry

TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN

FILE REFERENCE: PF-0615 US

CURRENT APPLICATION NUMBER: US/09/172,977

CURRENT FILING DATE: 1998-10-14

NUMBER OF SEQ ID NOS: 4

SOFTWARE: PERL Program

SEQ ID NO 4

LENGTH: 1839

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE: -

OTHER INFORMATION: g29491

US-09-172-977-4

Query Match 15.1%; Score 202.5; DB 2; Length 1839;

Best Local Similarity 29.3%; Pred. No. 2.8e-11;

Matches 56; Conservative 43; Mismatches 81; Indels 11; Gaps 4;

GENERAL INFORMATION: US/09172977

Patent No. 5989863

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Guegler, Karl J.

APPLICANT: Corley, Neil C.

APPLICANT: Yue, Henry

TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN

FILE REFERENCE: PF-0615 US

CURRENT APPLICATION NUMBER: US/09/172,977

CURRENT FILING DATE: 1998-10-14

NUMBER OF SEQ ID NOS: 4

SOFTWARE: PERL Program

SEQ ID NO 4

LENGTH: 1839

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE: -

OTHER INFORMATION: g29491

US-09-172-977-4

QY 55 DASVSSPQAGSSSLKHSTTLTNRQRGNEVSALPATLDSLSIHQLAAQGELDQLKEHLRKG 114

Db 5 DAAQKS-DSGEKFNSSQRRKPKKSDSNA-----SFLRAARAGNLDKWEYLKGGI 55

QY 115 NLVKNKPDGERGFTPLIWASAFGEIETVRFLEWAGDPHILAKERESALSASTGGYTDIVG 174

Db 56 D-INTCNQNGNLALHLAAKEGHVGLVQELLGRGSSVDSATKKGNHTALHIASLAGQAEVVK 114

QY 175 LLERDVVDINIYDNGGTPLLYAVRGNHVKVEALLARGADLTTEADSGYTPMDLAVAG 234

Db 115 VLVKEGANINAQSQNGFTPLYMAAQENHIDVVKYLLENGANQSTATEDGFTPLAVALQQG 174

QY 235 YRK-VQQVIEN 244

Db 175 HNQAVALLEN 185

RESULT 13

US-09-404-108-4

Query Match 14.7%; Score 197; DB 2; Length 843;

Best Local Similarity 32.4%; Pred. No. 3e-11;

Matches 48; Conservative 35; Mismatches 63; Indels 2; Gaps 2;

GENERAL INFORMATION: US/09404108

Patent No. 6590077

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Guegler, Karl J.

APPLICANT: Corley, Neil C.

APPLICANT: Yue, Henry

TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN

FILE REFERENCE: PF-0615 US

CURRENT APPLICATION NUMBER: US/09/404,108

CURRENT FILING DATE: 1999-09-23

EARLIER APPLICATION NUMBER: 09/172,977

EARLIER FILING DATE: 1998-10-14

NUMBER OF SEQ ID NOS: 4

SOFTWARE: PERL Program

SEQ ID NO 4

LENGTH: 1839

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE: -

OTHER INFORMATION: g29491

US-09-172-977-4

QY 55 DASVSSPQAGSSSLKHSTTLTNRQRGNEVSALPATLDSLSIHQLAAQGELDQLKEHLRKG 114

Db 5 DAAQKS-DSGEKFNSSQRRKPKKSDSNA-----SFLRAARAGNLDKWEYLKGGI 55

QY 115 NLVKNKPDGERGFTPLIWASAFGEIETVRFLEWAGDPHILAKERESALSASTGGYTDIVG 174

Db 56 D-INTCNQNGNLALHLAAKEGHVGLVQELLGRGSSVDSATKKGNHTALHIASLAGQAEVVK 114

QY 175 LLERDVVDINIYDNGGTPLLYAVRGNHVKVEALLARGADLTTEADSGYTPMDLAVAG 234

Db 115 VLVKEGANINAQSQNGFTPLYMAAQENHIDVVKYLLENGANQSTATEDGFTPLAVALQQG 174

QY 235 YRK-VQQVIEN 244

Db 175 HNQAVALLEN 185

RESULT 14

US-09-172-977-3

Query Match 14.7%; Score 197; DB 2; Length 843;

Best Local Similarity 32.4%; Pred. No. 3e-11;

Matches 48; Conservative 35; Mismatches 63; Indels 2; Gaps 2;

GENERAL INFORMATION: US/09172977

Patent No. 5989863

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Guegler, Karl J.

APPLICANT: Corley, Neil C.

APPLICANT: Yue, Henry

TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN

FILE REFERENCE: PF-0615 US

CURRENT APPLICATION NUMBER: US/09/172,977

CURRENT FILING DATE: 1998-10-14

NUMBER OF SEQ ID NOS: 4

SOFTWARE: PERL Program

SEQ ID NO 3

LENGTH: 843

TYPE: PRT

ORGANISM: Rattus norvegicus

FEATURE: -

OTHER INFORMATION: g1841966

US-09-172-977-3

QY 98 AAQGELDQLKEHLRKGDNLVNKPDERGFTPLIWASAFGEIETVRFLEWAGDPHILAKER 157

Db 4 ARAGNLDKWEYLKGGID-INTCNQNGNLALHLAAKEGHVGLVQELLGRGSSVDSATKKG 62

QY 158 ESALSASTGGYTDIVGLLERDVVDINIYDNGGTPLLYAVRGNHVKVEALLARGADLT 217

Db 63 NTALHIASLAGQAEVVKLVKEGANINAQSQNGFTPLYMAAQENHIDVVKYLLENGANQ 122

QY 218 TEADSGYTPMDLAVAGYRK-VQQVIEN 244

Db 123 TATEDGFTPLAVALQQGHNAVALLEN 150

RESULT 15

US-09-404-108-3

Query Match 14.7%; Score 197; DB 2; Length 843;

Best Local Similarity 32.4%; Pred. No. 3e-11;

Matches 48; Conservative 35; Mismatches 63; Indels 2; Gaps 2;

GENERAL INFORMATION: US/09404108

Patent No. 6590077

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Guegler, Karl J.

APPLICANT: Corley, Neil C.

APPLICANT: Yue, Henry

TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN

FILE REFERENCE: PF-0615 US

CURRENT APPLICATION NUMBER: US/09/404,108

CURRENT FILING DATE: 1999-09-23

EARLIER APPLICATION NUMBER: 09/172,977

EARLIER FILING DATE: 1998-10-14

NUMBER OF SEQ ID NOS: 4

SOFTWARE: PERL Program

SEQ ID NO 4

LENGTH: 1839

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE: -

OTHER INFORMATION: g29491

US-09-172-977-4

QY 55 DASVSSPQAGSSSLKHSTTLTNRQRGNEVSALPATLDSLSIHQLAAQGELDQLKEHLRKG 114

Db 5 DAAQKS-DSGEKFNSSQRRKPKKSDSNA-----SFLRAARAGNLDKWEYLKGGI 55

QY 115 NLVKNKPDGERGFTPLIWASAFGEIETVRFLEWAGDPHILAKERESALSASTGGYTDIVG 174

Db 56 D-INTCNQNGNLALHLAAKEGHVGLVQELLGRGSSVDSATKKGNHTALHIASLAGQAEVVK 114

QY 175 LLERDVVDINIYDNGGTPLLYAVRGNHVKVEALLARGADLTTEADSGYTPMDLAVAG 234

Db 115 VLVKEGANINAQSQNGFTPLYMAAQENHIDVVKYLLENGANQSTATEDGFTPLAVALQQG 174

QY 235 YRK-VQQVIEN 244

Db 175 HNQAVALLEN 185

RESULT 15

US-09-404-108-3

Query Match 14.7%; Score 197; DB 2; Length 843;

Best Local Similarity 32.4%; Pred. No. 3e-11;

Matches 48; Conservative 35; Mismatches 63; Indels 2; Gaps 2;

GENERAL INFORMATION: US/09404108

Patent No. 6590077

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Guegler, Karl J.

APPLICANT: Corley, Neil C.

APPLICANT: Yue, Henry

TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN

FILE REFERENCE: PF-0615 US

CURRENT APPLICATION NUMBER: US/09/404,108

CURRENT FILING DATE: 1999-09-23

EARLIER APPLICATION NUMBER: 09/172,977

EARLIER FILING DATE: 1998-10-14

NUMBER OF SEQ ID NOS: 4

SOFTWARE: PERL Program

SEQ ID NO 4

LENGTH: 1839

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE: -

OTHER INFORMATION: g29491

US-09-172-977-4

OTHER INFORMATION: g29491

US-09-404-108-4

Query Match 15.1%; Score 202.5; DB 4; Length 1839;

Best Local Similarity 29.3%; Pred. No. 2.8e-11;

Matches 56; Conservative 43; Mismatches 81; Indels 11; Gaps 4;

QY 55 DASVSSPQAGSSSLKHSTTLTNRQRGNEVSALPATLDSLSIHQLAAQGELDQLKEHLRKG 114

Db 5 DAAQKS-DSGEKFNSSQRRKPKKSDSNA-----SFLRAARAGNLDKWEYLKGGI 55

QY 115 NLVKNKPDGERGFTPLIWASAFGEIETVRFLEWAGDPHILAKERESALSASTGGYTDIVG 174

Db 56 D-INTCNQNGNLALHLAAKEGHVGLVQELLGRGSSVDSATKKGNHTALHIASLAGQAEVVK 114

QY 175 LLERDVVDINIYDNGGTPLLYAVRGNHVKVEALLARGADLTTEADSGYTPMDLAVAG 234

Db 115 VLVKEGANINAQSQNGFTPLYMAAQENHIDVVKYLLENGANQSTATEDGFTPLAVALQQG 174

QY 235 YRK-VQQVIEN 244

Db 175 HNQAVALLEN 185

RESULT 14

US-09-172-977-3

Query Match 14.7%; Score 197; DB 2; Length 843;

Best Local Similarity 32.4%; Pred. No. 3e-11;

Matches 48; Conservative 35; Mismatches 63; Indels 2; Gaps 2;

GENERAL INFORMATION: US/09172977

Patent No. 5989863

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Guegler, Karl J.

APPLICANT: Corley, Neil C.

APPLICANT: Yue, Henry

TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN

FILE REFERENCE: PF-0615 US

CURRENT APPLICATION NUMBER: US/09/172,977

CURRENT FILING DATE: 1998-10-14

NUMBER OF SEQ ID NOS: 4

SOFTWARE: PERL Program

SEQ ID NO 3

LENGTH: 843

TYPE: PRT

ORGANISM: Rattus norvegicus

FEATURE: -

OTHER INFORMATION: g1841966

US-09-172-977-3

QY 98 AAQGELDQLKEHLRKGDNLVNKPDERGFTPLIWASAFGEIETVRFLEWAGDPHILAKER 157

Db 4 ARAGNLDKWEYLKGGID-INTCNQNGNLALHLAAKEGHVGLVQELLGRGSSVDSATKKG 62

QY 158 ESALSASTGGYTDIVGLLERDVVDINIYDNGGTPLLYAVRGNHVKVEALLARGADLT 217

Db 63 NTALHIASLAGQAEVVKLVKEGANINAQSQNGFTPLYMAAQENHIDVVKYLLENGANQ 122

QY 218 TEADSGYTPMDLAVAGYRK-VQQVIEN 244

Db 123 TATEDGFTPLAVALQQGHNAVALLEN 150

RESULT 15

US-09-404-108-3

Query Match 14.7%; Score 197; DB 2; Length 843;

Best Local Similarity 32.4%; Pred. No. 3e-11;

Matches 48; Conservative 35; Mismatches 63; Indels 2; Gaps 2;

GENERAL INFORMATION: US/09404108

Patent No. 6590077

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Guegler, Karl J.

APPLICANT: Corley, Neil C.

APPLICANT: Yue, Henry

TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN

FILE REFERENCE: PF-0615 US

CURRENT APPLICATION NUMBER: US/09/404,108

CURRENT FILING DATE: 1999-09-23

EARLIER APPLICATION NUMBER: 09/172,977

EARLIER FILING DATE: 1998-10-14

NUMBER OF SEQ ID NOS: 4

SOFTWARE: PERL Program

SEQ ID NO 4

LENGTH: 1839

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE: -

OTHER INFORMATION: g29491

US-09-172-977-4

Search completed: April 29, 2004, 06:27:53
Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 29, 2004, 06:16:22 ; Search time 60 Seconds
(without alignments)
1224.373 Million cell updates/sec

Title: US-09-840-243C-11
Perfect score: 1341
Sequence: 1 MELTQPAEDLIQTQQTBASE.....VIENHILKLFQSNLVPADPE 260

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1341	100.0	260	3 AAY59539	Aay59539 Human ank
2	1341	100.0	260	3 AAY79411	Aay79411 Human MHC
3	1341	100.0	260	7 ADD27859	Add27859 Human ank
4	1113.5	83.0	269	3 AAY59590	Aay59590 Ankyrin r
5	625	46.6	229	5 ABP41881	Abp41881 Human ova
6	621.5	46.3	313	4 AAG66309	Aag66309 Human ank
7	621.5	46.3	313	4 AAB94322	Aab94322 Human pro
8	621.5	46.3	314	7 ADE60905	Ade60905 Human pro
9	621.5	46.3	314	7 ADE60909	Ade60909 Human pro
10	612.5	45.7	312	7 ADD47704	Add47704 Rat Prote
11	607.5	45.3	263	4 AAU20646	Aau20646 Human sec
12	495	36.9	152	4 ABB50161	Abb50161 Human tra
13	421.5	31.4	84	3 AAG01584	Aag01584 Human sec
14	393	29.3	119	4 AAU20665	Aau20665 Human sec
15	364	27.1	105	4 AAU20558	Aau20558 Human sec
16	299	22.3	81	5 ABB97342	Abb97342 Novel hum
17	274	20.4	234	4 ABB61859	Abb61859 Drosophil
18	257	19.2	49	4 AAM21225	Aam21225 Peptide #
19	257	19.2	49	4 ABB43550	Abb43550 Peptide #
20	257	19.2	49	5 ABG46308	Abg46308 Human pep
21	216.5	16.1	747	6 ADA54356	Ada54356 Human pro
22	216.5	16.1	1429	6 AAE36185	Aae36185 Human CGD
23	212	15.8	705	4 AAG75604	Aag75604 Human col
24	212	15.8	1184	7 ADC06846	Adc06846 Human kid
25	212	15.8	1762	5 AAU96841	Aau96841 Rat kidin

26	212	15.8	1762	7 ADC06848	Adc06848 Rat Kidin
27	212	15.8	1771	7 ADC06847	Adc06847 Human Kid
28	210.5	15.7	2443	4 ABB60521	Abb60521 Drosophil
29	207.5	15.5	166	5 ABB78585	Abb78585 3 ankyrin
30	207.5	15.5	1142	7 ADC06850	Adc06850 Kidins 22
31	207.5	15.5	1715	4 AAM39025	Aam39025 Human pol
32	207.5	15.5	1715	4 AAM38993	Aam38993 Human pol
33	207.5	15.5	1715	5 AAE25144	Aae25144 Human ARM
34	207.5	15.5	1715	5 AAE25143	Aae25143 Rat ARMS
35	207.5	15.5	1715	5 AAU96840	Aau96840 Human kid
36	207.5	15.5	1715	6 AAE32128	Aae32128 Human cyt
37	207.5	15.5	1715	7 ADC06849	Adc06849 Rat Kidin
38	207.5	15.5	1753	6 ADA09888	Ada09888 Human rec
39	206.5	15.4	1763	5 AAU80244	Aau80244 Rat kidin
40	205.5	15.3	342	4 ABB59641	Abb59641 Drosophil
41	205	15.3	348	2 AAW76775	Aaw76775 D. immitti
42	205	15.3	348	2 AAW70607	Aaw70607 Ankyrin p
43	205	15.3	348	3 AAB11588	Aab11588 D. immitti
44	205	15.3	348	5 AAC21367	Aao21367 Ankyrin p
45	205	15.3	1745	2 AAW76776	Aaw76776 D. immitti

ALIGNMENTS

RESULT 1

AAAY59539
ID AAAY59539 standard; protein; 260 AA.

AC AAAY59539;

XX
DT 03-APR-2000 (first entry)

XX
DE Human ankyrin family protein, ANFP.

XX
KW Human; ankyrin family protein; ANFP; autoimmune disorder; inflammation;
KW atherosclerosis; inflammatory disorder; proliferative disorder; AIDS;
KW vesicle-trafficking disorder; allergy; amyloidosis; anaemia; asthma;
KW bronchitis; Crohn's disease; atopic dermatitis; diabetes mellitus;
KW irritable bowel syndrome; osteoporosis; rheumatoid arthritis; cirrhosis;
KW hepatitis; ulcerative colitis; cancer; hypercholesterolaemia; therapy;
KW diagnosis.

OS Homo sapiens.

XX
PN US5989863-A.

XX
PD 23-NOV-1999.

XX
PF 14-OCT-1998; 98US-00172977.

XX
PR 14-OCT-1998; 98US-00172977.

XX
PA (INCY-) INCYTE PHARM INC.

XX
PI Tang YT, Corley NC, Yue H, Guegler KJ;

XX
DR WPI; 2000-095634/08.

XX
DR N-PSDB; AAZ49052.

XX
PT Polynucleotide sequence encoding a human ankyrin family protein useful
for diagnosis or treatment of autoimmune, inflammatory, proliferative and
vesicle-trafficking disorders.

XX
PS Claim 1; Fig 1; 34pp; English.

XX
CC This sequence is the human ankyrin family protein, ANFP, of the
invention. Host cells containing an expression vector containing the
polynucleotide sequence can be cultured to produce ANFP, which can be
used for diagnosis or treatment of autoimmune, inflammatory,
proliferative and vesicle-trafficking disorders. Disorders which can be
treated include acquired immune deficiency syndrome (AIDS), allergies,
amyloidosis, anaemia, asthma, atherosclerosis, bronchitis, Crohn's

CC disease, atopic dermatitis, diabetes mellitus, irritable bowel syndrome,
CC myocardial or pericardial inflammation, osteoporosis, rheumatoid
CC arthritis, cirrhosis, hepatitis, ulcerative colitis, cancer and
CC hypercholesterolaemia. The polynucleotide sequences can also be used as a
CC hybridisation probe to detect ANFP-encoding polynucleotides or to screen
CC samples. Purified ANFP can be used to produce antibodies or to screen
CC libraries of pharmaceutical agents to find agents that specifically bind
CC ANFP. The DNA and its antisense sequence can be used in therapeutic
CC compositions e.g. to regulate gene function. The DNA sequence can be used
CC for diagnostic purposes to detect and quantitate gene expression in
CC biopsied tissues and to indicate the absence, presence and excess
CC expression of ANFP and monitor its levels during therapeutic intervention
XX
SQ Sequence 260 AA;

Query Match 100.0%; Score 1341; DB 3; Length 260;
Best Local Similarity 100.0%; Pred. No. 5.9e-128;
Matches 260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MELTQPAEDLIQTQTPASELGDPEDEEADGSDTVVLSLFPCTPEPVNPEPDASVSS 60
Db 1 MELTQPAEDLIQTQTPASELGDPEDEEADGSDTVVLSLFPCTPEPVNPEPDASVSS 60
QY 61 PQAGSSLKHSHTLTNRQRGNEVSALPATLDSLSIHQLAAQGEILDQLKEHLRKGDNLVNKP 120
Db 61 PQAGSSLKHSHTLTNRQRGNEVSALPATLDSLSIHQLAAQGEILDQLKEHLRKGDNLVNKP 120
QY 121 DERGFTPLIWASAFGEIETVRFLLEWGADPHILAKERESALSASTGGYTDIVGLLLERD 180
Db 121 DERGFTPLIWASAFGEIETVRFLLEWGADPHILAKERESALSASTGGYTDIVGLLLERD 180
QY 181 VDINIYDWNGGTPLLYAVRGNHVKCVALLARGADLTTEADSGYTPMDLAVALGYRKVQ 240
Db 181 VDINIYDWNGGTPLLYAVRGNHVKCVALLARGADLTTEADSGYTPMDLAVALGYRKVQ 240
QY 241 VIENHILKLFQSNLVPADPE 260
Db 241 VIENHILKLFQSNLVPADPE 260

RESULT 2
AA79411
ID AA79411 standard; protein; 260 AA.
XX AA79411;
AC AA79411;
XX
DT 01-AUG-2000 (first entry)
DE Human MHC class II gene transcription factor RFXANK.
XX
KW RFXANK; HsRFXANK; human; transcription factor; MHC class II;
KW chromosome 19p12; immunosuppressive; immunomodulator; antiinflammatory;
KW antidiabetic; antiarthritic; therapy; inflammation; autoimmune diseases;
KW transplant rejection; insulin dependent diabetes; multiple sclerosis;
KW lupus erythematosus; rheumatoid arthritis; immunodeficiency.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 122..156
FT /note= "ankyrin repeat region 1"
FT Region 157..188
FT /note= "ankyrin repeat region 2"
FT Region 189..222
FT /note= "ankyrin repeat region 3"
XX
PN EP995798-A1.
XX
PD 26-APR-2000.
XX
PF 24-OCT-1998; 98EP-00120085.
XX
PR 24-OCT-1998; 98EP-00120085.

XX (NOVI-) NOVIMMUNE SA.
PA Masternak K, Reith W, Mach B;
PI WPI; 2000-294958/26.
XX N-PSDB; AAZ94868.
DR Novel isolated transcription factor, RFXANK, useful for treating MHC
DR class II deficiency and autoimmune disorders, e.g. insulin dependent
XX diabetes and multiple sclerosis, restores the functional transcription of
XX MHC class II genes.
PS Claim 1; Fig 3; 48pp; English.
XX
CC The present sequence is that of human RFXANK, a novel transcription
CC factor that is a subunit of the RFX heterotrimeric transcription complex
CC that binds to the conserved X box motif of all MHC class II gene
CC promoters. The RFXANK gene is mutated in complementation group B MHC II
CC deficiency patients. Mutations identified in patients include aberrant
CC splicing and short deletions in exon 6. The nvention provides inhibitors
CC of RFXANK including antibodies, single chain antibodies, dominant
CC negative mutants, antisense molecules and ribozymes. The inhibitors may
CC be used in therapy or prevention of diseases associated with aberrant
CC expression of MHC class II genes and/or as an immunosuppressive agents,
CC e.g. to treat inflammation, insulin dependent diabetes, multiple sclerosis,
CC transplanted organs, autoimmune diseases or rejection of
CC lupus erythematosus and rheumatoid arthritis. The compositions may also
CC be used to treat the autosomal recessive disease MHC class II deficiency.
CC Since RFXANK does not play any other major role in the transcriptional
CC control of genes other than MHC class II genes, its inhibitors are devoid
CC of other undesirable inhibitory effects
XX
SQ Sequence 260 AA;

Query Match 100.0%; Score 1341; DB 3; Length 260;
Best Local Similarity 100.0%; Pred. No. 5.9e-128;
Matches 260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MELTQPAEDLIQTQTPASELGDPEDEEADGSDTVVLSLFPCTPEPVNPEPDASVSS 60
Db 1 MELTQPAEDLIQTQTPASELGDPEDEEADGSDTVVLSLFPCTPEPVNPEPDASVSS 60
QY 61 PQAGSSLKHSHTLTNRQRGNEVSALPATLDSLSIHQLAAQGEILDQLKEHLRKGDNLVNKP 120
Db 61 PQAGSSLKHSHTLTNRQRGNEVSALPATLDSLSIHQLAAQGEILDQLKEHLRKGDNLVNKP 120
QY 121 DERGFTPLIWASAFGEIETVRFLLEWGADPHILAKERESALSASTGGYTDIVGLLLERD 180
Db 121 DERGFTPLIWASAFGEIETVRFLLEWGADPHILAKERESALSASTGGYTDIVGLLLERD 180
QY 181 VDINIYDWNGGTPLLYAVRGNHVKCVALLARGADLTTEADSGYTPMDLAVALGYRKVQ 240
Db 181 VDINIYDWNGGTPLLYAVRGNHVKCVALLARGADLTTEADSGYTPMDLAVALGYRKVQ 240
QY 241 VIENHILKLFQSNLVPADPE 260
Db 241 VIENHILKLFQSNLVPADPE 260

RESULT 3
ADD27859
ID ADD27859 standard; protein; 260 AA.
XX
AC ADD27859;
XX
DT 15-JAN-2004 (first entry)
XX Human ankyrin family protein (ANFP).
DE Human; ankyrin family protein; ANFP; cell growth;
XX altered cell morphology; NIH3T3 cell; autoimmune disorder;
KW inflammatory disorder; cell proliferative disorder;

QY 231 VALGYRKVQVQVNIENHILKLFQSNLVPADPE 260
DB 240 VALGYRKVQVQVMESHILRLFQSTLGPVDPE 269

RESULT 5
ABP41881
ID ABP41881 standard; protein; 229 AA.
XX
AC ABP41881;
XX
DT 22-AUG-2002 (first entry)
XX
DE Human ovarian antigen HHK22, SEQ ID NO:3013.
XX
KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
KW inflammatory condition; immune disorder; blood disorder;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disorder; urinary system disorder; drug screening;
KW gene therapy; chromosome mapping; forensic analysis;
KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW antiinflammatory; gynaecological; reproductive.
XX
OS Homo sapiens.
XX
PN WO200200677-A1.
XX
PD 03-JAN-2002.
PF 07-JUN-2001; 2001WO-US018569.
XX
PR 07-JUN-2000; 2000US-0209467P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Birse CE, Rosen CA;
XX
DR WPI; 2002-147878/19.
DR N-PSDB; ABQ54958.
XX
PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
PT cancer), immune disorders, cardiovascular disorders and neurological
PT diseases.
XX
PS Claim 11; SEQ ID NO 3013; 2922pp; English.
XX
CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
CC encompasses polypeptides 90% identical and polynucleotides 95% identical
CC to the sequences of the invention. The invention additionally relates to
CC recombinant vectors and host cells comprising human ovarian antigen
CC polynucleotides, antibodies against human ovarian antigens, and the use
CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
CC treating, prognosing or preventing various ovary and/or breast-related
CC disorders. Such conditions include ovarian cancer and breast cancer, and
CC metastatic tumours of ovarian or breast origin, reproductive system
CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
CC vaginitis), immune disorders (e.g., congenital and acquired
CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
CC respiratory disorders, neurological disorders, gastrointestinal disorders
CC and urinary system disorders. Ovarian antigen polypeptides and
CC polynucleotides may also be used in screening for compounds which
CC modulate ovarian antigen expression or activity. The polynucleotides may
CC further be used for gene therapy, chromosome mapping, in the
CC identification of individuals and in forensic analysis, and the

CC polypeptides may be used as food additives or to prepare antibodies
CC useful in disease diagnosis, drug targeting and phenotyping. The present
CC sequence represents a human ovarian antigen of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 229 AA;

Query Match 46.6%; Score 625; DB 5; Length 229;
Best Local Similarity 58.6%; Pred. No. 4.9e-55;
Matches 130; Conservative 26; Mismatches 58; Indels 8; Gaps 3;

QY 37 TVVLSLPPCTPEPVNPEPDASV-----SSPQACSSSLKHSTTLTNRQGNVSALPATLD 90
DB 7 TTFVHLAECNIH-TSPSPGIQVRHVYTPSTTKHFSPKQSTTLTNKHRGNVSTPPLLAN 65
QY 91 SLSIHQLAAQGGELDQKEHLRKGDNLVNKPDERGFTPLIWASAFGEIETVRELLWEGADP 150
DB 66 SLSVHQLAAQGGEMLYLATRIEQ-ENVINHTDEGFTPLMWAAGQIAVVEFLQNGADP 124
QY 151 HILAKERESALSLASTGGYTDIVGLLLERDVINIDWNGGTPLLYAVRGNHVKCVKVEALL 210
DB 125 QLLGKGRESALSLACSGYTDIVKMLLDGVDVNEYDWWNGGTPLLYAVHGNHVKCVKMLL 184
QY 211 ARGADLTTEADSGYTPMDLAVLGYRKVQVNIENHILKLFQS 252
DB 185 ESGADPTIETDSGYNMMDLAVLGYRSVQQVIESHLKLLQN 226

RESULT 6
AAG66309

ID AAG66309 standard; protein; 313 AA.

XX
AC AAG66309;

XX
DT 09-OCT-2001 (first entry)

XX
DE Human ankyrin-like protein 34..

XX
KW Human; ankyrin-like protein 34; cytostatic; virucidal; immunomodulatory;
KW antiinflammatory; haemostatic; gene therapy; malignant tumour;
KW haemopathy; HIV infection; immunological disease; inflammation.

XX
OS Homo sapiens.

XX
PN WO200155194-A1.

XX
PD 02-AUG-2001.

XX
PF 21-JAN-2001; 2001WO-CN000085.

XX
PR 28-JAN-2000; 2000CN-00111595.

XX
PA (BIOD-) BIODOOR GENE TECHNOLOGY LTD SHANGHAI.

XX
PI Mao Y, Xie Y;

XX
DR WPI; 2001-483222/52.

XX
DR N-PSDB; AAH75702.

XX
PT New human ankyrin-like protein 34 for diagnosing and treating malignant
PT tumor, hemopathy, human immunodeficiency virus infection, immunological
PT diseases and various inflammations.

XX
PS Claim 1; Page 31-32; 38pp; Chinese.

XX
CC The present sequence is the protein sequence for human ankyrin-like
CC protein 34. The ankyrin-like protein and its coding sequence are useful
CC in the diagnosis and treatment of malignant tumour, haemopathy, HIV
CC infection, immunological diseases and various inflammations

XX
SQ Sequence 313 AA;

Query Match 46.3%; Score 621.5; DB 4; Length 313;
Best Local Similarity 60.6%; Pred. No. 1.8e-54;
Matches 126; Conservative 26; Mismatches 49; Indels 7; Gaps 2;
QY 51 NPEPDASV-----SSPQAGSSSLKHSTTLTNRQRGNEVSALPATLDSLSIHQLAAQGGELD 104
Db 104 SPSPGIQVRHVYTPSTTKHFSPKQSTTLTNKHRGNEVSTTPLLANSLSVHQLAAQGGEML 163
QY 105 QLKEHLRKGDNLVNKPDERGFTPLIWSAFGEIETVRFLEWGADPHILAKERESALSILA 164
Db 164 YLATRIEQ-ENVINHTDEEGFTPLMWAHAHQIAVVEFLQNGADPQLLGKGRESALSILA 222
QY 165 STGGYTDIVGLLLERDVNINIDWNGGTPLLYAVRGHNHVKVEALLARGADLTTEADSGY 224
Db 223 CSKGYTDIVKMLLDGVDVNEVDWNGGTPLLYAVHGNHVKCVKMLLESGADPTIETDSGY 282
QY 225 TPMDLAVALGYRKVQVQVNIENHILKLFQS 252
Db 283 NSMDLAVALGYRSVQVQVIESHLLKLLQN 310

RESULT 7
AAB94322
ID AAB94322 standard; protein; 313 AA.
AC AAB94322;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:14803.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX

Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
XX
PS Claim 8; SEQ ID NO 14803; 2537pp + Sequence Listing; English.
XX
CC The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of

the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention

XX SQ Sequence 313 AA;

Query Match 46.3%; Score 621.5; DB 4; Length 313;
Best Local Similarity 60.6%; Pred. No. 1.8e-54;
Matches 126; Conservative 26; Mismatches 49; Indels 7; Gaps 2;

QY 51 NPEPDASV-----SSPQAGSSSLKHSTTLTNRQRGNEVSALPATLDSLSIHQLAAQGGELD 104
Db 104 SPSPGIQVRHVYTPSTTKHFSPKQSTTLTNKHRGNEVSTTPLLANSLSVHQLAAQGGEML 163
QY 105 QLKEHLRKGDNLVNKPDERGFTPLIWSAFGEIETVRFLEWGADPHILAKERESALSILA 164
Db 164 YLATRIEQ-ENVINHTDEEGFTPLMWAHAHQIAVVEFLQNGADPQLLGKGRESALSILA 222
QY 165 STGGYTDIVGLLLERDVNINIDWNGGTPLLYAVRGHNHVKVEALLARGADLTTEADSGY 224
Db 223 CSKGYTDIVKMLLDGVDVNEVDWNGGTPLLYAVHGNHVKCVKMLLESGADPTIETDSGY 282
QY 225 TPMDLAVALGYRKVQVQVNIENHILKLFQS 252
Db 283 NSMDLAVALGYRSVQVQVIESHLLKLLQN 310

RESULT 8
ADE60905
ID ADE60905 standard; protein; 314 AA.
XX
AC ADE60905;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human Protein AK022876, SEQ ID NO 6819.
XX
KW Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
OS Homo sapiens.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
DR WPI; 2003-268312/26.
DR GENBANK; AK022876.
XX
PT New composition comprising two or more isolated polypeptides, useful for preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 314 AA;

Query Match 46.3%; Score 621.5; DB 7; Length 314;
Best Local Similarity 60.6%; Pred. No. 1.8e-54;
Matches 126; Conservative 26; Mismatches 49; Indels 7; Gaps 2;
QY 51 NPEPDASV-----SSPQAGSSSLKHSTTLTNRQRGNEVSALPATLDSLSIHQLAAQGEGLD 104
Db 104 SPSPGIQVRHVYTPSTTKHFSPIKQSTTLTNKRGNEVSTTPLLANSLSVHQLAAQGEML 163
QY 105 QLKEHLRKGDNLVNKPDERGFTPLIWASAFGEIETVRFLEWGDADPHILAKERESALSLSLA 164
Db 164 YLATRIEQ-ENVINHTDEEGFTPLMWAHAHQIAVVEFLQNGADPQLLGKRESALSLSLA 222
QY 165 STGGYTDIVGLLLERDVIDINIDYDNGGTPLLYAVRGNHVKVCVEALLARGADLTTEADSGY 224
Db 223 CSKGYTDIVKMLDCGVDVNEYDNGGTPLLYAVHGNHVKVCVMLESADPTIETDSGY 282
QY 225 TPMDLAVALGYRKVQVQVNIENHILKLFQS 252
Db 283 NSMDLAVALGYRSVQVQVIESHLLKLLQN 310

RESULT 9
ADE60909
ID ADE60909 standard; protein; 314 AA.
XX ADE60909;
XX 29-JAN-2004 (first entry)
XX Human Protein AK022876, SEQ ID NO 6823.
XX Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX Homo sapiens.
XX WO2003016475-A2.
XX 27-FEB-2003.
XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX Woolf C, D'urso D, Befort K, Costigan M;
PI WPI; 2003-268312/26.
XX GENBANK; AK022876.
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX Claim 1; Page; 1017pp; English.

CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 314 AA;

Query Match 46.3%; Score 621.5; DB 7; Length 314;
Best Local Similarity 60.6%; Pred. No. 1.8e-54;
Matches 126; Conservative 26; Mismatches 49; Indels 7; Gaps 2;
QY 51 NPEPDASV-----SSPQAGSSSLKHSTTLTNRQRGNEVSALPATLDSLSIHQLAAQGEGLD 104
Db 104 SPSPGIQVRHVYTPSTTKHFSPIKQSTTLTNKRGNEVSTTPLLANSLSVHQLAAQGEML 163
QY 105 QLKEHLRKGDNLVNKPDERGFTPLIWASAFGEIETVRFLEWGDADPHILAKERESALSLSLA 164
Db 164 YLATRIEQ-ENVINHTDEEGFTPLMWAHAHQIAVVEFLQNGADPQLLGKRESALSLSLA 222
QY 165 STGGYTDIVGLLLERDVIDINIDYDNGGTPLLYAVRGNHVKVCVEALLARGADLTTEADSGY 224
Db 223 CSKGYTDIVKMLDCGVDVNEYDNGGTPLLYAVHGNHVKVCVMLESADPTIETDSGY 282
QY 225 TPMDLAVALGYRKVQVQVNIENHILKLFQS 252
Db 283 NSMDLAVALGYRSVQVQVIESHLLKLLQN 310

RESULT 10
ADD47704
ID ADD47704 standard; protein; 312 AA.
XX

PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.

PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-02559678P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-451931/48.

XX N-PSDB; AAS33355.

PT New nucleic acids and polypeptides, useful for diagnosing, preventing or treating medical conditions.

XX Claim 11; SEQ ID NO 638; 753pp; English.

CC The invention relates to novel isolated nucleic acid molecules (I) encoding human secreted proteins (II). (I) and (II) are used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. (I) and (II) may be used in the prevention, treatment and diagnosis of diseases associated with inappropriate expression of secreted proteins. (I) and complementary sequences may also be used as DNA probes in diagnostic assays (e.g. polymerase chain reactions (PCR)) to detect and quantitate the presence of similar nucleic acid sequences in samples, and so which patients may be in need of restorative therapy. (II) may also be used as antigens in the production of antibodies and in assays to identify modulators (agonists and antagonists) of the expression and activity of the secreted proteins. The anti-(II) antibodies and antagonists may also be used to down regulate expression and activity of (II). The anti-(II) antibodies may also be used as diagnostic agents for detecting the presence of (II) in samples (e.g. by enzyme linked immunosorbant assay (ELISA)). The disorders include for example: immune/autoimmune diseases (e.g. HIV (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis and multiple sclerosis), cancers and hyperproliferative disorders (e.g. melanomas, neoplasms of the breast or liver, Sezary syndrome and Gaucher's disease), neurological diseases (e.g. Alzheimer's disease, Parkinson's disease and Charcot-Marie-Tooth disease), cardio-/cerebrovascular disorders (e.g. cardiac arrest, tachycardia, angina and thrombosis), infections caused by bacteria, viruses and fungi and agonists (e.g. corneal infections). (I) and (II), agonists, antagonists and antibodies can also be used to promote wound healing, maintain organs and before transplantation, and support cell culture of primary tissues.

Query Match 45.3%; Score 607.5; DB 4; Length 263;
Best Local Similarity 59.4%; Pred. No. 3.7e-53;
Matches 123; Conservative 26; Mismatches 51; Indels 7; Gaps 2;

```

QY 51 NPEPDASV-----SSPQAGSSSLKHSTLTNRQRGNEVSALPATLDSLSIHQLAAQGELD 104
Db 36 SPSPGIQVRHVYTPSTTKHFSPIKQSTLTNTKHRGNEVSTTPLLANSLSVHQLAAQGEML 95

QY 105 QLKEHLRKGDNLVKNPDERGFTPLIWASAFGEIETVRFLLEWGADPHILAKERESALSIA 164
Db 96 YLATRIEQ-ENVINHTDEEGFTPLMWAHAHQIAVVEFLIQNGADPQLLGKRESALSIA 154

QY 165 STGGYTDIVGLLLERDVIDINIYDWNCGTPLLAYVRGNHVKCV EALLARGADITTEADSGY 224
Db 155 CSKGYTDIVKMLLDCGV DNV EYDWNCGTPLLAYAVHGNHKKCVKMLLES GADPTIETDSGY 214

QY 225 TPMDLAVALGYRKVQVQVIENHILKLFQ 251
Db 215 NSMDLAVALGYRSVQQVIESHLNCPK 241

```

RESULT 12
ABB50161
ID ABB50161 standard; protein: 152 AA.

DT 05-FEB-2002 (first entry)

Human transcription factor TRFX-12.

Human; transcription factor; TRFX; cell proliferative disease;
 autoimmune disease; inflammation; neurological disease;
 developmental disorder; cancer; AIDS; infection; cytostatic; anti-HIV;
 neuroprotective; antiinflammatory; gene therapy.

OS Homo sapiens.

PN WO200172777-A2.

PD 04-OCT-2001.

PF 13-MAR-2001; 2001WO-US008117.

PR 13-MAR-2000; 2000US-0188986P.

PA (INCY-) INCYTE GENOMICS INC.

PI Hillman JL, Baughn MR, Yue H, Lal P, Lu DAM, Patterson C;
PI Azimzai Y, Bandman O, Tang YT, Mathur P, Shah P, Au-Young J;
PI Reddy R;

DR WPI; 2001-570896/64.

DR N-PSDB; ABA82985.

PT Novel transcription factor polypeptides, used to treat diseases
PT associated with altered activity and expression of TRFX, and to screen
PT for agents capable of modulating its activity.

PS Claim 1; Page 151-152; 327pp; English.

The present sequence is the protein sequence for a human transcription factor. The transcription factor and its coding sequence are useful in the diagnosis, treatment and prevention of diseases associated with altered expression of the transcription factor e.g. cell proliferative, autoimmune/inflammatory, neurological and developmental disorders. A number of specific disorders/diseases are given in the specification, including: arteriosclerosis, cirrhosis, hepatitis, cancers, AIDS, allergies, anaemia, asthma, autoimmune thyroiditis, bronchitis, atopic dermatitis, diabetes mellitus, emphysema, Goodpasture's syndrome, gout, Grave's disease, multiple sclerosis, osteoarthritis, pancreatitis, psoriasis, rheumatoid arthritis, systemic lupus erythematosus, ulcerative colitis, uveitis, Alzheimer's disease, Huntington's disease, Parkinson's disease, stroke, and viral, bacterial, fungal and protozoal infections.

Sequence 152 AA;

		Query Match	36.9%;	Score 495;	DB 4;	Length 152;	
		Best Local Similarity	68.3%;	Pred. NO. 4.8e-42;			
		Matches 95;	Conservative 16;	Mismatches 28;	Indels 0;	Gaps 0;	
QY	114	DNLVKNKPDERTPLIWASAFGEIETVRFLLEWGADPHILAKERESALSLASTGGYTDIV	173				
		: : : : : : : : : :					
Db	11	ENVINHTDEGFTPLMWAHAHQIAVVEFLLQNGADPQLLGKGRESALSLACSKGYTDIV	70				
		:					
QY	174	GLLERDVDINIYDWNGGTPLL YAVRGNHV KVCVEALLARGADLTTEADSGYTPMDLAVAL	233				
		: : : : : : : : : : :					
Db	71	KMLDCGV DVNEYDWNGGTPLL YAVHGNHV KVCVKMLLES GADPTIETDSGYNSMDLAVAL	130				
		:					
QY	234	GYRKVQQVTIENHILKL FQS	252				
		: : : : :					
Db	131	GYRSVOOVIESHL LKLLON	149				
		:					

RESULT 13
AAG01584
ID AAG01584 standard; protein: 84 AA.

DT 06-OCT-2000 (first entry)

Human secreted protein, SEO ID NO: 5665.

Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping.

OS Homo sapiens.

PN EP1033401-A2.

06-SEP-2000.

21-FEB-2000: 2000EP-00200610

PR 26-FEB-1999; 99US-0122487P.

PA (GEST) GENSET.

PI Dumas Milne Edwards J, Duclert A, Giordano J;

DR WPI: 2000-500381/45.

DR N-PSDB; AAC01590.

New nucleic acid that is a 5' expressed sequence tag (5' EST) for PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for PT diagnostic, forensic, gene therapy and chromosome mapping procedures.

PS Claim 13: SEQ ID NO 5665; 71pp + Sequence Listing; English.

The present sequence is a polypeptide encoded by one of a large number of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs derived from 30 different tissues. EST sequences usually correspond mainly to the 3' untranslated region (UTR) of the mRNA because they are often obtained from oligo-dT primed cDNA libraries. Such ESTs are not well suited for isolating cDNA sequences derived from the 5' ends of mRNAs and even in those cases where longer cDNA sequences have been obtained, the full 5' UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and chromosome mapping procedures. They are used to obtain upstream regulatory sequences and to design expression and secretion vectors.

Sequence 84 AA:

Query Match	31.4%;	Score 421.5;	DB 3;	Length 84;
Best Local Similarity	98.8%;	Pred. No. 6.2e-35;		
Matches 84;	Conservative	0;	Mismatches 0;	Indels 1;
				Gaps 1;

Oy 1 MELTQPAEDLIQTQTPASELGDPEDPGEEAAGSDTVVLSLFPCTPEPVNPEPDASVSS 60
Db 1 MELTQPAEDLIQTQTPASELGDPEDPGEEAAGSDTVVLSLFPCTPEPVNPEPDASVSS 60
Oy 61 PQAGSSLKHSHTLTNRQRGNEVSAL 85
Db 61 PQ-GSSLKHSHTLTNRQRGNEVSAL 84

RESULT 14

AAU20665

ID AAU20665 standard; protein; 119 AA.

XX AC AAU20665;

XX DT 04-DEC-2001 (first entry)

XX DE Human secreted protein, Seq ID No 657.

XX KW Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;

XX KW rheumatoid arthritis; antiarteriosclerotic; cardiac; vascular;

XX KW cerebroprotective; thrombolytic; antimicrobial; ophthalmological;

XX KW cytotostatic; Alzheimer's disease; Parkinson's disease; human; cancer;

XX KW multiple sclerosis; cancer; hyperproliferative disorder; infection;

XX KW Gaucher's disease; neurological disease; cerebrovascular disorder;

XX KW thrombosis; wound healing.

XX OS Homo sapiens.

XX PN WO200155326-A2.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001347.

XX PR 31-JAN-2000; 2000US-0179065P.

XX PR 04-FEB-2000; 2000US-0180628P.

XX PR 24-FEB-2000; 2000US-0184664P.

XX PR 02-MAR-2000; 2000US-0186350P.

XX PR 16-MAR-2000; 2000US-0189874P.

XX PR 17-MAR-2000; 2000US-0190076P.

XX PR 18-APR-2000; 2000US-0198123P.

XX PR 19-MAY-2000; 2000US-0205515P.

XX PR 07-JUN-2000; 2000US-0209467P.

XX PR 28-JUN-2000; 2000US-0214886P.

XX PR 30-JUN-2000; 2000US-0215135P.

XX PR 07-JUL-2000; 2000US-0216647P.

XX PR 11-JUL-2000; 2000US-0217487P.

XX PR 11-JUL-2000; 2000US-0217496P.

XX PR 14-JUL-2000; 2000US-0218290P.

XX PR 26-JUL-2000; 2000US-0220963P.

XX PR 14-AUG-2000; 2000US-0224518P.

XX PR 14-AUG-2000; 2000US-0224519P.

XX PR 14-AUG-2000; 2000US-0225213P.

XX PR 14-AUG-2000; 2000US-0225266P.

XX PR 14-AUG-2000; 2000US-0225267P.

XX PR 14-AUG-2000; 2000US-0225268P.

XX PR 14-AUG-2000; 2000US-0225270P.

XX PR 14-AUG-2000; 2000US-0225447P.

XX PR 14-AUG-2000; 2000US-0225757P.

XX PR 14-AUG-2000; 2000US-0225758P.

XX PR 14-AUG-2000; 2000US-0225759P.

XX PR 18-AUG-2000; 2000US-0226279P.

XX PR 22-AUG-2000; 2000US-0226681P.

XX PR 22-AUG-2000; 2000US-0226688P.

XX PR 22-AUG-2000; 2000US-0227182P.

XX PR 23-AUG-2000; 2000US-0227009P.

XX PR 30-AUG-2000; 2000US-0228924P.

XX PR 01-SEP-2000; 2000US-0229287P.

XX PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.

PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-451931/48.
DR N-PSDB; AAS33374.
XX
PT New nucleic acids and polypeptides, useful for diagnosing, preventing or
XX treating medical conditions.
PS Claim 11; SEQ ID NO 657; 753pp; English.
XX
CC The invention relates to novel isolated nucleic acid molecules (I)
CC encoding human secreted proteins (II). (I) and (II) are used to prevent,
CC treat or ameliorate a medical condition in e.g. humans, mice, rabbits,
CC goats, horses, cats, dogs, chickens or sheep. (I) and (II) may be used in
CC the prevention, treatment and diagnosis of diseases associated with
CC inappropriate expression of secreted proteins. (I) and complementary
CC sequences may also be used as DNA probes in diagnostic assays (e.g.
CC polymerase chain reactions (PCR)) to detect and quantitate the presence
CC of similar nucleic acid sequences in samples, and so which patients may
CC be in need of restorative therapy. (II) may also be used as antigens in
CC the production of antibodies and in assays to identify modulators
CC (agonists and antagonists) of the expression and activity of the secreted
CC proteins. The anti-(II) antibodies and antagonists may also be used to
CC down regulate expression and activity of (II). The anti-(II) antibodies
CC may also be used as diagnostic agents for detecting the presence of (II)
CC in samples (e.g. by enzyme linked immunosorbant assay (ELISA)). The
CC disorders include for example: immune/autoimmune diseases (e.g. HIV
CC (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis
CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.
CC melanomas, neoplasms of the breast or liver, Sezary syndrome and
CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,
CC Parkinson's disease and Charcot-Marie-Tooth disease), cardio-/
CC cerebrovascular disorders (e.g. cardiac arrest, tachycardia, angina and
CC thrombosis), infections caused by bacteria, viruses and fungi and ocular
CC disorders (e.g. corneal infections). (I) and (II), agonists, antagonists
CC and antibodies can also be used to promote wound healing, maintain organs
CC before transplantation, and support cell culture of primary tissues.

Query Match 29.3%; Score 393; DB 4; Length 119;
Best Local Similarity 72.6%; Pred. No. 8.3e-32;
Matches 77; Conservative 9; Mismatches 20; Indels 0; Gaps 0;

QY 147 GADPHILAKRESALSASTGGYTDIVGLLLERDVIDINIYDNGGTPLLYAVRGNHVKCV 206
DB 11 GADPQLLGKRESALSACSKGYTDIVKMLDCCGVDDVNEYDNGGTPLLYAVHGNHVKCV 70

QY 207 EALLARGADLTTEADSGYTPMDLAVALGYRKVQVQVIENHILKLFQS 252
Db 71 KMLLESGADPTIETDSGYNMMDLAVALGYRSVQVQVIESHLKLQN 116
RESULT 15
AAU20558
ID AAU20558 standard; protein; 105 AA.
XX AC AAU20558;
XX DT 04-DEC-2001 (first entry)
XX DE Human secreted protein, Seq ID No 550.
XX KW Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;
KW rheumatoid arthritis; antiarteriosclerotic; cardiant; vascular;
KW cerebroprotective; thrombolytic; antimicrobial; ophthalmological;
KW cytostatic; Alzheimer's disease; Parkinson's disease; human; cancer;
KW multiple sclerosis; cancer; hyperproliferative disorder; infection;
KW Gaucher's disease; neurological disease; cerebrovascular disorder;
KW thrombosis; wound healing.
XX OS Homo sapiens.
XX PN WO200155326-A2.
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US001347.
XX PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.

Db 65 DPTIETDSGYNSMDLAVALGYRSVQQVIESHLLKLLQN 102

Search completed: April 29, 2004, 06:25:49
Job time : 63 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

QM protein - protein search, using sw model

Run on: April 29, 2004, 06:26:53 ; Search time 47 Seconds
(without alignments)
1533.389 Million cell updates/sec

Title: US-09-840-243C-11
Perfect score: 1341
Sequence: 1 MELTQPAEDLIQTQTPASE.....VIENHILKLFQSNLVPADPE 260

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1138120 seqs, 277189581 residues

Total number of hits satisfying chosen parameters: 1138120

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1341	100.0	260	9 US-09-840-243B-11	Sequence 11, Appl
2	1341	100.0	260	9 US-09-840-243B-12	Sequence 12, Appl
3	1113.5	83.0	269	9 US-09-840-243B-13	Sequence 13, Appl
4	625	46.6	229	15 US-10-264-049-3013	Sequence 3013, Ap
5	561.5	41.9	220	9 US-09-840-243B-18	Sequence 18, Appl
6	542.5	40.5	218	9 US-09-840-243B-19	Sequence 19, Appl
7	495	36.9	152	12 US-10-221-625-12	Sequence 12, Appl
8	257	19.2	49	9 US-09-864-761-47859	Sequence 47859, A
9	216.5	16.1	747	15 US-10-094-749-1924	Sequence 1924, Ap
10	212	15.8	705	14 US-10-106-698-6378	Sequence 6378, Ap
11	212	15.8	1184	14 US-10-117-229-3	Sequence 3, Appli
12	212	15.8	1184	14 US-10-117-229-9	Sequence 9, Appli
13	212	15.8	1762	14 US-10-117-229-7	Sequence 7, Appli
14	212	15.8	1771	14 US-10-117-229-4	Sequence 4, Appli
15	212	15.8	1771	14 US-10-117-229-11	Sequence 11, Appl

16	207.5	15.5	1715	14	US-10-021-571-2	Sequence 2, Appli
17	207.5	15.5	1715	14	US-10-021-571-4	Sequence 4, Appli
18	207.5	15.5	1715	14	US-10-117-229-2	Sequence 2, Appli
19	207.5	15.5	1715	14	US-10-117-229-6	Sequence 6, Appli
20	202.5	15.1	720	16	US-10-433-794-20	Sequence 20, Appl
21	202.5	15.1	765	12	US-10-182-243-56	Sequence 56, Appl
22	200.5	15.0	301	12	US-10-112-944-738	Sequence 738, App
23	200.5	15.0	301	12	US-10-112-944-739	Sequence 739, App
24	200.5	15.0	301	12	US-10-112-944-740	Sequence 740, App
25	199.5	14.9	657	15	US-10-104-047-2529	Sequence 2529, Ap
26	199	14.8	1498	14	US-10-021-571-8	Sequence 8, Appli
27	198.5	14.8	765	14	US-10-128-174-3	Sequence 3, Appli
28	198.5	14.8	765	14	US-10-128-174-34	Sequence 34, Appl
29	198.5	14.8	765	14	US-10-128-174-35	Sequence 35, Appl
30	198.5	14.8	765	14	US-10-128-174-36	Sequence 36, Appl
31	198.5	14.8	765	14	US-10-128-174-37	Sequence 37, Appl
32	198.5	14.8	765	14	US-10-128-174-38	Sequence 38, Appl
33	198.5	14.8	765	14	US-10-128-174-39	Sequence 39, Appl
34	198.5	14.8	765	14	US-10-128-174-40	Sequence 40, Appl
35	198.5	14.8	765	14	US-10-128-174-41	Sequence 41, Appl
36	198.5	14.8	765	14	US-10-128-174-42	Sequence 42, Appl
37	198.5	14.8	765	14	US-10-128-174-43	Sequence 43, Appl
38	198.5	14.8	765	14	US-10-128-174-44	Sequence 44, Appl
39	198.5	14.8	1330	15	US-10-108-260A-3237	Sequence 3237, Ap
40	196.5	14.7	551	9	US-09-835-788A-17	Sequence 17, Appl
41	196.5	14.7	551	14	US-10-175-042-17	Sequence 17, Appl
42	192	14.3	740	9	US-09-835-788A-12	Sequence 12, Appl
43	192	14.3	740	14	US-10-175-042-12	Sequence 12, Appl
44	191	14.2	426	9	US-09-908-711-70	Sequence 70, Appl
45	190	14.2	309	12	US-10-425-114-5427	Sequence 5427, A

ALIGNMENTS

RESULT 1

US-09-840-243B-11
; Sequence 11, Application US/09840243B
; Patent No. US20020156258A1
; GENERAL INFORMATION:
; APPLICANT: MASTERNAK, Krzysztof
; APPLICANT: REITH, Walter
; APPLICANT: MACH, Bernard
; TITLE OF INVENTION: New Transcription Factor of MHC Class II Genes, Substances
; TITLE OF INVENTION: Capable of Inhibiting This New Transcription Factor and
; TITLE OF INVENTION: Medical Uses of These Substances
; FILE REFERENCE: 010830-117
; CURRENT APPLICATION NUMBER: US/09/840,243B
; CURRENT FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: EP 98120085.0
; PRIOR FILING DATE: 1998-10-24
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-840-243B-11

Query Match	100.0%;	Score 1341;	DB 9;	Length 260;
Best Local Similarity	100.0%;	Pred. No. 1.1e-115;		
Matches 260;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MELTQPAEDLIQTQTPASELGDPGEEAAGSDTIVLSLFPCTPEPVNPEPDASVSS	60	
Db	1	MELTQPAEDLIQTQTPASELGDPGEEAAGSDTIVLSLFPCTPEPVNPEPDASVSS	60	
QY	61	POAGSSLKHSHTLTNRQRGNEVSALPATLDSLSIHQLAAQAGELDQLKEHLRGDNLVNKP	120	
Db	61	POAGSSLKHSHTLTNRQRGNEVSALPATLDSLSIHQLAAQAGELDQLKEHLRGDNLVNKP	120	
QY	121	DERGFTPLIWASAFGEIETVRFLLEWGADPHILAKERESALSASTGGYTDIVGLLLERD	180	

Db 121 DERGFTPLIWASAFGEIETVRFLEWGADPHILAKERESALSASTGGYTDIVGLLLERD 180
QY 181 VDINIYDNGGTPLLYAVRGNHVKCVEALLARGADLTTEADSGYTPMDLAVALGYRKVQQ 240
Db 181 VDINIYDNGGTPLLYAVRGNHVKCVEALLARGADLTTEADSGYTPMDLAVALGYRKVQQ 240
QY 241 VIENHILKLFQSNLVPADPE 260
Db 241 VIENHILKLFQSNLVPADPE 260

RESULT 2
US-09-840-243B-12
; Sequence 12, Application US/09840243B
; Patent No. US20020156258A1
; GENERAL INFORMATION:
; APPLICANT: MASTERNAK, Krzysztof
; APPLICANT: REITH, Walter
; APPLICANT: MACH, Bernard
; TITLE OF INVENTION: New Transcription Factor of MHC Class II Genes, Substances
; TITLE OF INVENTION: Capable of Inhibiting This New Transcription Factor and
; TITLE OF INVENTION: Medical Uses of These Substances
; FILE REFERENCE: 010830-117
; CURRENT APPLICATION NUMBER: US/09/840,243B
; CURRENT FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: EP 98120085.0
; PRIOR FILING DATE: 1998-10-24
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-840-243B-12

Query Match 100.0%; Score 1341; DB 9; Length 260;
Best Local Similarity 100.0%; Pred. No. 1.1e-115;
Matches 260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MELTQPAEDLIQTQTPASELGDPEDEAAGSDTVVLSLFPCTPEPVNPEPDASVSS 60
Db 1 MELTQPAEDLIQTQTPASELGDPEDEAAGSDTVVLSLFPCTPEPVNPEPDASVSS 60
QY 61 PQAGSSLKHSHTLTNRQNEVSALPATLDSLSIHQLAAQGELDQKHEHLRKGDNLVNKP 120
Db 61 PQAGSSLKHSHTLTNRQNEVSALPATLDSLSIHQLAAQGELDQKHEHLRKGDNLVNKP 120
QY 121 DERGFTPLIWASAFGEIETVRFLEWGADPHILAKERESALSASTGGYTDIVGLLLERD 180
Db 121 DERGFTPLIWASAFGEIETVRFLEWGADPHILAKERESALSASTGGYTDIVGLLLERD 180
QY 181 VDINIYDNGGTPLLYAVRGNHVKCVEALLARGADLTTEADSGYTPMDLAVALGYRKVQQ 240
Db 181 VDINIYDNGGTPLLYAVRGNHVKCVEALLARGADLTTEADSGYTPMDLAVALGYRKVQQ 240
QY 241 VIENHILKLFQSNLVPADPE 260
Db 241 VIENHILKLFQSNLVPADPE 260

RESULT 3
US-09-840-243B-13
; Sequence 13, Application US/09840243B
; Patent No. US20020156258A1
; GENERAL INFORMATION:
; APPLICANT: MASTERNAK, Krzysztof
; APPLICANT: REITH, Walter
; APPLICANT: MACH, Bernard
; TITLE OF INVENTION: New Transcription Factor of MHC Class II Genes, Substances
; TITLE OF INVENTION: Capable of Inhibiting This New Transcription Factor and
; TITLE OF INVENTION: Medical Uses of These Substances
; FILE REFERENCE: 010830-117
; CURRENT APPLICATION NUMBER: US/09/840,243B

; CURRENT FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: EP 98120085.0
; PRIOR FILING DATE: 1998-10-24
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-09-840-243B-13

Query Match 83.0%; Score 1113.5; DB 9; Length 269;
Best Local Similarity 81.9%; Pred. No. 1.3e-94;
Matches 221; Conservative 15; Mismatches 23; Indels 11; Gaps 2;
QY 1 MELTQPAEDLIQTQTPASELGDPEDEAAGSDTVVLSLFPCTPEPVNPEPDASVSS 60
Db 1 MEPTQVAENLVNQQPPVDPDLEDPEDTRDESPENDTVVLSLFPCTPDVNPEDASASS 60
QY 61 PQAGSSLKHSHTLTNRQNEVSALPATLDSLSIHQLAAQGELDQKHEHLRKGDNLVNKP 112
Db 61 LQ-GSFLKHSHTLTNRQNEVSALPATLDSLSIHQLAAQGELDQKHEHLRKGACPACTC 119
QY 113 --GDNLVNKPDERGFTPLIWASAFGEIETVRFLEWGADPHILAKERESALSASTGGYT 170
Db 120 LSGNNLINKPDERGFTPLIWASAFGEIETVRFLLDWGADPHILAKERESALSASTGGYT 179
QY 171 DIVGLLLERDVIDINIYDNGGTPLLYAVRGNHVKCVEALLARGADLTTEADSGYTPMDLA 230
Db 180 DIVRLLDRDVIDINIYDNGGTPLLYAVRGNHVKCVEALLARGADLTTEADSGYTPMDLA 239
QY 231 VALGYRKVQQVIENHILKLFQSNLVPADPE 260
Db 240 VALGYRKVQQVMESHILRLFQSTLGPVDPE 269

RESULT 4
US-10-264-049-3013
; Sequence 3013, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 3013
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-049-3013

Query Match 46.6%; Score 625; DB 15; Length 229;
Best Local Similarity 58.6%; Pred. No. 1.7e-49;
Matches 130; Conservative 26; Mismatches 58; Indels 8; Gaps 3;
QY 37 TVVLSLFPCTPEPVNPEPDASV-----SSPQAGSSLKHSHTLTNRQNEVSALPATLD 90
Db 7 TPFVFLAECNIH-TSPSPGIQVRHVYTPSTTKHFSPIKQSTLTNKHGRGNEVSTPILLAN 65
QY 91 SLSIHQLAAQGELDQKHEHLRKGDNLVNKPDERGFTPLIWASAFGEIETVRFLEWGADP 150
Db 66 SLSVHQLAAQGEMLYLAIRIEQ-ENVINHTDEEGFTPLMWAHAHQIAVVEFLQNGADP 124
QY 151 HILAKERESALSASTGGYTDIVGLLLERDVIDINIYDNGGTPLLYAVRGNHVKCVEALL 210
Db 125 QLLGKGRESALSACSGYTDIVKMLLDGVDVNEYDNGGTPLLYAVHGNHVKCVKMLL 184

Search completed: April 29, 2004, 06:34:56
Job time : 48 secs

QY 220 AD SGYTPMDLAVAGY-RKVQQVIENH 245
Db 200 GANSM TALIVAVKGGYTQSVKEILKRN 226

RESULT 14

US-10-117-229-4
; Sequence 4, Application US/10117229
; Publication No. US20030190625A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc.
; TITLE OF INVENTION: Human Kidins220Pc
; FILE REFERENCE: 9U 104 R1
; CURRENT APPLICATION NUMBER: US/10/117,229
; CURRENT FILING DATE: 2002-04-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 1771
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-117-229-4

Query Match 15.8%; Score 212; DB 14; Length 1771;
Best Local Similarity 34.7%; Pred. No. 5.3e-10;
Matches 51; Conservative 33; Mismatches 61; Indels 2; Gaps 2;

QY 100 QGELDQLKEHLRKGDNLVNKPDGERGFTPLIWASAFGEIETVRFLEWGADPHILAKERES 159
Db 81 EGHVHIVEELLKCGVNLEHR-DMGGWTALMWACYKGRTDVVELLSHGANPSVTGLYSVY 139
QY 160 ALSLASTGGYTDIVGLLLERDV DINIYDWNGGTPLLYAVRGNHVKVEALLARGADLTTE 219
Db 140 PIIWAAGRGHADIVHLLQLONGAKVNCSDKYGTTPLVWAARKGHLECVKHLLAMGADV DQE 199

QY 220 AD SGYTPMDLAVAGY-RKVQQVIENH 245
Db 200 GANSM TALIVAVKGGYTQSVKEILKRN 226

RESULT 15

US-10-117-229-11
; Sequence 11, Application US/10117229
; Publication No. US20030190625A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc.
; TITLE OF INVENTION: Human Kidins220Pc
; FILE REFERENCE: 9U 104 R1
; CURRENT APPLICATION NUMBER: US/10/117,229
; CURRENT FILING DATE: 2002-04-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 1771
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-117-229-11

Query Match 15.8%; Score 212; DB 14; Length 1771;
Best Local Similarity 34.7%; Pred. No. 5.3e-10;
Matches 51; Conservative 33; Mismatches 61; Indels 2; Gaps 2;

QY 100 QGELDQLKEHLRKGDNLVNKPDGERGFTPLIWASAFGEIETVRFLEWGADPHILAKERES 159
Db 81 EGHVHIVEELLKCGVNLEHR-DMGGWTALMWACYKGRTDVVELLSHGANPSVTGLYSVY 139
QY 160 ALSLASTGGYTDIVGLLLERDV DINIYDWNGGTPLLYAVRGNHVKVEALLARGADLTTE 219
Db 140 PIIWAAGRGHADIVHLLQLONGAKVNCSDKYGTTPLVWAARKGHLECVKHLLAMGADV DQE 199

QY 220 AD SGYTPMDLAVAGY-RKVQQVIENH 245
Db 200 GANSM TALIVAVKGGYTQSVKEILKRN 226

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 29, 2004, 06:16:22 ; Search time 17 Seconds
(without alignments)
796.367 Million cell updates/sec

Title: US-09-840-243C-11
Perfect score: 1341
Sequence: 1 MELTQPAEDLIQTQTPASE.....VIENHILKLFQSNLVPADPE 260

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues
Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1341	100.0	260	1	RFXK HUMAN
2	1113.5	83.0	269	1	RFXK_MOUSE
3	621.5	46.3	313	1	ANRA_HUMAN
4	612.5	45.7	312	1	ANRA_MOUSE
5	206.5	15.4	1267	1	HMT1_HUMAN
6	202.5	15.1	3924	1	ANK2_HUMAN
7	193	14.4	4377	1	ANK3_HUMAN
8	189	14.1	1210	1	BAT8_HUMAN
9	188	14.0	1263	1	BAT8_MOUSE
10	187	13.9	475	1	GASZ_MOUSE
11	187	13.9	1327	1	TNK1_HUMAN
12	184	13.7	1880	1	ANK1_HUMAN
13	183	13.6	1862	1	ANK1_MOUSE
14	182.5	13.6	231	1	PSDA_MOUSE
15	181.5	13.5	768	1	YB23_HUMAN
16	180	13.4	475	1	GASZ_PANTR
17	179	13.3	475	1	GASZ_HUMAN
18	176.5	13.2	226	1	PSDA_HUMAN
19	176.5	13.2	231	1	PSDA_RAT
20	175	13.0	328	1	ANR2_MOUSE
21	175	13.0	435	1	AKR_ARATH
22	174.5	13.0	360	1	Y045_METMA
23	173	12.9	333	1	ANR2_HUMAN
24	173	12.9	415	1	ANRX_MOUSE
25	173	12.9	1150	1	ANS1_MOUSE
26	172.5	12.9	1059	1	Y379_HUMAN
27	171.5	12.8	420	1	ANRX_HUMAN
28	169.5	12.6	451	1	ILK_CAVPO
29	169.5	12.6	741	1	RN5A_HUMAN
30	169	12.6	256	1	YI61_PYRAE
31	168	12.5	776	1	ANR5_HUMAN
32	166.5	12.4	1442	1	DAK1_MOUSE
33	165.5	12.3	452	1	ILK1_HUMAN

34	165.5	12.3	452	1	ILK2_HUMAN	P57043	homo sapien
35	164	12.2	642	1	YA2A_SCHPO	Q09701	schizosacch
36	163.5	12.2	518	1	ASB3_HUMAN	Q9Y575	homo sapien
37	163	12.2	117	1	MTPN_MOUSE	P80144	mus musculus
38	163	12.2	118	1	MTPN_CHICK	Q91955	gallus gall
39	163	12.2	1401	1	LATA_LATWA	P23631	latrodectus
40	162.5	12.1	1134	1	ANS1_HUMAN	Q92625	homo sapien
41	162	12.1	656	1	FEM1_CAEEL	P17221	caenorhabdi
42	161	12.0	117	1	MTPN_HUMAN	P58546	homo sapien
43	161	12.0	1432	1	DAK1_HUMAN	P53355	homo sapien
44	160.5	12.0	452	1	ILK_MOUSE	O55222	mus musculus
45	160.5	12.0	525	1	ASB3_MOUSE	Q9WV72	mus musculus

ALIGNMENTS

RESULT 1
RFXK_HUMAN
ID RFXK_HUMAN STANDARD; PRT; 260 AA.
AC O14593; O95839;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE DNA-binding protein RFXANK (Regulatory factor X subunit B) (RFX-B)
DE (Ankyrin repeat family A protein 1).
GN RFXANK OR RFXB OR ANKRA1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND MASS SPECTROMETRY.
RC TISSUE=B-cell;
RX MEDLINE=99021383; PubMed=9806546;
RA Masternak K., Barras E., Zufferey M., Conrad B., Corthals G.,
RA Aebersold R., Sanchez J.-C., Hochstrasser D.F., Mach B., Reith W.;
RT "A gene encoding a novel RFX-associated transactivator is mutated in
RT the majority of MHC class II deficiency patients.";
RL Nat. Genet. 20:273-277(1998).
RN [2]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 79-95; 180-198; 100-210 AND
RP 238-248.
RC TISSUE=Lymphoblast;
RX MEDLINE=99170284; PubMed=10072068;
RA Nagarajan U.M., Louis-Plence P., Desandro A., Nilsen R., Bushey A.,
RA Boss J.M.;
RT "RFX-B is the gene responsible for the most common cause of the bare
RT lymphocyte syndrome, an MHC class II immunodeficiency.";
RL Immunity 10:153-162(1999).
RN [3]
RP ERRATUM.
RA Nagarajan U.M., Louis-Plence P., Desandro A., Nilsen R., Bushey A.,
RA Boss J.M.;
RL Immunity 10:399-399(1999).
RN [4]
RP SEQUENCE FROM N.A.
RA Zhou J., Guan Z., Gu J., Ye M., Fu G., Zhang Q., Xu S., He K.,
RA Chen S., Mao M., Chen Z.;
RT "Hematopoiesis-derived ankyrin-like gene.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 65-260 FROM N.A.
RA Lamerdin J.E., McCready P.M., Adamson A.W., Burkhardt-Schultz K.,
RA Christensen M., Gordon L., Kyle A., Ramirez M., Stilwagen S.,
RA Garnes J., Danganan L., Bruce R., Quan G., Montgomery M., Ow D.,
RA Kobayashi A., Olsen A.O., Carrano A.V.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [6]
RP VARIANT BLS II GROUP B PRO-195.
RX MEDLINE=20192025; PubMed=10725724;
RA Nagarajan U.M., Peijnenburg A., Gobin S.J., Boss J.M.,
RA van den Elsen P.J.;

RT "Novel mutations within the RFX-B gene and partial rescue of MHC and
RT related genes through exogenous class II transactivator in
RT RFX-B-deficient cells."
RL J. Immunol. 164:3666-3674(2000).
CC -!- FUNCTION: Activates transcription from class II MHC promoters.
CC Activation requires the activity of the MHC class II
CC transactivator (MHC2TA). May regulate other genes in the cell. RFX
CC binds the X1 box of MHC-II promoters. Isoform RFX-B-DELTA5 is not
CC involved in the positive regulation of MHC class II genes.
CC -!- SUBUNIT: Rfx consists of at least three different subunits; RFXAP,
CC RFX5 and RFX-B/RFXANK; with each subunit representing a separate
CC complementation group. RFX forms cooperative DNA binding complexes
CC with X2BP and CBF/NF-Y. RFX associates with MHC2TA to form an
CC active transcriptional complex.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=014593-1; Sequence=Displayed;
CC Name=RFX-B-delta5;
CC IsoId=014593-2; Sequence=VSP_000283, VSP_000284;
CC -!- TISSUE SPECIFICITY: Ubiquitous.
CC -!- DOMAIN: The third ankyrin repeat is required for association with
CC the two other RFX subunits; RFX5 and RFXAP.
CC -!- DISEASE: Defects in RFXANK are a cause of bare lymphocyte syndrome
CC type II (BLS II) [MIM:209920]; also known as hereditary MHC class
CC II deficiency or HLA class II-deficient combined immunodeficiency.
CC BLS II is a form of severe combined immunodeficiency disease
CC (SCID) characterized by a profound defect in constitutive and
CC interferon-gamma induced MHC II expression, an absence of cellular
CC and humoral T-cell response to antigen challenge,
CC hypogammaglobulinemia and impaired antibody production. The
CC consequence include extreme susceptibility to viral, bacterial and
CC fungal infections. RFXANK is linked with BLS II complementation
CC group B.
CC -!- SIMILARITY: Contains 5 ANK repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AF094760; AAC69883.1; -.
DR EMBL; AF105427; AAD17972.1; -.
DR EMBL; AF105428; AAD17973.1; -.
DR EMBL; AF077196; AAD26991.1; -.
DR EMBL; AC003110; AAB86654.1; -.
DR HSSP; P42773; 1BU9.
DR TRANSFAC; T05441; -.
DR TRANSFAC; T05442; -.
DR Genew; HGNC:9987; RFXANK.
DR MIM; 603200; -.
DR MIM; 209920; -.
DR GO; GO:0003677; F:DNA binding; TAS.
DR GO; GO:0003712; F:transcription cofactor activity; TAS.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ank; 4.
DR SMART; SM00248; ANK; 4.
DR PROSITE; PS50088; ANK_REPEAT; 3.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
KW DNA-binding; Transcription regulation; Activator; Nuclear protein;
KW Repeat; ANK repeat; Alternative splicing; Phosphorylation;
KW Disease mutation; SCID; Polymorphism.
FT REPEAT 89 118 ANK 1.
FT REPEAT 123 152 ANK 2.
FT REPEAT 156 185 ANK 3.
FT REPEAT 189 218 ANK 4.
FT REPEAT 222 251 ANK 5.
FT VARSPLIC 63 63 Missing (in isoform RFX-B-delta5).
FT /FTid=VSP_000283.

FT VARSPLIC 91 113 SLSIHQLAAQGGELDQLKEHLRKG -> C (in isoform
FT RFX-B-delta5).
FT /FTid=VSP_000284.
FT VARIANT 195 195 L->P (in BLS II group B).
FT /FTid=VAR_009941.
FT VARIANT 251 251 Q->E (in dbSNP:1802498).
FT /FTid=VAR_014472.
SQ SEQUENCE 260 AA; 28102 MW; 6280B490F54816D2 CRC64;
Query Match 100.0%; Score 1341; DB 1; Length 260;
Best Local Similarity 100.0%; Pred.No. 1.5e-97;
Matches 260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MELTQPAEDLIQTQTPASELGDPEDEEADGSDTVVLSLFPCTPEPVNPEPDASVSS 60
DB 1 MELTQPAEDLIQTQTPASELGDPEDEEADGSDTVVLSLFPCTPEPVNPEPDASVSS 60
QY 61 PQAGSSLKHSSTLTNRQRGNEVSALPATLDSLSIHQLAAQGGELDQLKEHLRKGDNLVNKP 120
DB 61 PQAGSSLKHSSTLTNRQRGNEVSALPATLDSLSIHQLAAQGGELDQLKEHLRKGDNLVNKP 120
QY 121 DERGFTPLIWSAFGEIETVRFLLEWGADPHILAKERESALSASTGGYTDIVGLLLERD 180
DB 121 DERGFTPLIWSAFGEIETVRFLLEWGADPHILAKERESALSASTGGYTDIVGLLLERD 180
QY 181 VDINIYDNGGTPLLYAVRGNHVKCVALLARGADLTTEADSGYTPMDLAVLGYRKVQQ 240
DB 181 VDINIYDNGGTPLLYAVRGNHVKCVALLARGADLTTEADSGYTPMDLAVLGYRKVQQ 240
QY 241 VIENHILKLFQSNLVPADPE 260
DB 241 VIENHILKLFQSNLVPADPE 260
RESULT 2
RFXK_MOUSE STANDARD; PRT; 269 AA.
AC Q9Z205;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, last annotation update)
DE DNA-binding protein RFXANK (Regulatory factor X subunit B) (RFX-B)
DE (Regulatory factor X-associated ankyrin-containing protein) (Ankyrin
DE repeat-containing adapter protein Tvl-1).
GN RFXANK OR RFXB OR TVL1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Spleen;
RX MEDLINE=99021383; PubMed=9806546;
RA Mastersnak K., Barras E., Zufferey M., Conrad B., Corthals G., Reith W.;
RA Aebersold R., Sanchez J.-C., Hochstrasser D.F., Mach B., Reith W.;
RT "A gene encoding a novel RFX-associated transactivator is mutated in
RT the majority of MHC class II deficiency patients."
RL Nat. Genet. 20:273-277(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=T-cell;
RX MEDLINE=99262619; PubMed=10329666;
RA Lin J.-H., Makris A., McMahon C., Bear S.E., Patriotis C.,
RA Prasad V.R., Brent R., Golemis E.A., Tschlis P.N.;
RT "The ankyrin repeat-containing adaptor protein tvl-1 is a novel
RT substrate and regulator of raf-1."
RL J. Biol. Chem. 274:14706-14715(1999).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM SHORT).
RC STRAIN=FVB/N; TISSUE=Salivary gland;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RL human and mouse cDNA sequences";
CC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: May facilitate endocytosis by linking megalin to
CC components of the cytoskeleton or endocytic machinery.
CC -!- SUBUNIT: Interacts directly and specifically with the cytoplasmic
CC tail of megalin (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic and associated with membranes
CC (By similarity).
CC -!- SIMILARITY: Contains 3 ANK repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF314032; AAK01621.1; -;
DR EMBL; AK022876; BAB14288.1; -;
DR EMBL; AF251051; AAK34941.1; -;
DR EMBL; BC012917; AAH12917.1; -;
DR HSSP; P80144; 2MYO.
DR Genew; HGNC:13208; ANKRA2.
DR MIM; 605787; -;
DR GO; GO:0005829; C:cytosol; IDA.
DR GO; GO:0016020; C:membrane; IDA.
DR GO; GO:0030169; F:low-density lipoprotein binding; IDA.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ank; 3.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 3.
DR PROSITE; PS50088; ANK_REPEAT; 3.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
KW Cytoskeleton; ANK repeat; Repeat.
FT REPEAT 181 213 ANK 1.
FT REPEAT 214 246 ANK 2.
FT REPEAT 247 279 ANK 3.
SQ SEQUENCE 313 AA; 34272 MW; 31C653B10B4ED6E1 CRC64;

Query Match 46.3%; Score 621.5; DB 1; Length 313;
Best Local Similarity 60.6%; Pred. No. 2.5e-41;
Matches 126; Conservative 26; Mismatches 49; Indels 7; Gaps 2;

QY 51 NPEPDASV-----SSPQAGSSSLKHSTLTNRQRGNEVSALPATLDSLSIHQLAAQGELD 104
Db 104 SPSPGIQVRHVYTPSTTKHFSPIKQSTLTNKHGRNEVSTTPLLANSLSVHQLAAQGEML 163

QY 105 QLKEHLRKGNLVNKPDERGFTPLIWASAFGEIETVRFLEWGADPHILAKERESALSLSLA 164
Db 164 YLATRIEQ-ENVINHTDEEGFTPLMWAHAHQIAVVEFLLQNGADPQLLGKGRESALSLSLA 222

QY 165 STGGYTDIVGLLLERDVVDINIYDWNNGTPLLAVRGNHVKVCVEALLARGADLTTEADSGY 224
Db 223 CSKGYTDIVKMLLDGVDVNEYDWNNGTPLLAVRGNHVKVCVLMLESADPTIETDGSY 282

QY 225 TMDLAVALGYRKVQVQVIENHILKLFQS 252
Db 283 NSMDLAVALGYRSVQVQVIESHLLKLLQN 310

RESULT 4
ANRA_MOUSE STANDARD; PRT; 312 AA.
ID ANRA_MOUSE
AC Q99PE2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ankyrin-repeat family A protein 2 (REFANK-like 2).
GN ANKRA2 OR ANKRA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., INTERACTION WITH MEGALIN, AND SUBCELLULAR
RP LOCATION.
RX MEDLINE=20547411; PubMed=11095640;
RA Rader K., Orlando R.A., Lou X., Farquhar M.G.;
RT "Characterization of ANKRA, a novel ankyrin repeat protein that
RT interacts with the cytoplasmic domain of megalin.";
RL J. Am. Soc. Nephrol. 11:2167-2178(2000).
CC -!- FUNCTION: May facilitate endocytosis by linking megalin to
CC components of the cytoskeleton or endocytic machinery.
CC -!- SUBUNIT: Interacts directly and specifically with the cytoplasmic
CC tail of megalin.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic and associated with membranes.
CC -!- SIMILARITY: Contains 3 ANK repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF314031; AAK01620.1; -;
DR HSSP; P80144; 2MYO.
DR MGD; MGI:1915808; Ankra2.
DR GO; GO:0005624; C:membrane fraction; IDA.
DR GO; GO:0005515; F:protein binding; IPI.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ank; 3.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 3.
DR PROSITE; PS50088; ANK_REPEAT; 3.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
KW Cytoskeleton; ANK repeat; Repeat.
FT REPEAT 180 212 ANK 1.
FT REPEAT 213 245 ANK 2.
FT REPEAT 246 278 ANK 3.
SQ SEQUENCE 312 AA; 34062 MW; 3B52B1415B7A5AFA CRC64;

Query Match 45.7%; Score 612.5; DB 1; Length 312;
Best Local Similarity 60.1%; Pred. No. 1.3e-40;
Matches 125; Conservative 25; Mismatches 51; Indels 7; Gaps 2;

QY 51 NPEPDASV-----SSPQAGSSSLKHSTLTNRQRGNEVSALPATLDSLSIHQLAAQGELD 104
Db 103 SPSPGIQVRHVYTPSTTKHFSPIKQSTLTNKHGRNEVSTTPLLANSLSAHQLAAQGEML 162

QY 105 QLKEHLRKGNLVNKPDERGFTPLIWASAFGEIETVRFLEWGADPHILAKERESALSLSLA 164
Db 163 YLATRIEQ-ENVINHTDEEGFTPLMWAHAHQIAVVEFLLQNGADPQLLGKGRESALSLSLA 221

QY 165 STGGYTDIVGLLLERDVVDINIYDWNNGTPLLAVRGNHVKVCVEALLARGADLTTEADSGY 224
Db 222 CSKGYTDIVKMLLDGVDVNEYDWNNGTPLLAVRGNHVKVCVLMLENGADPTIETDGSY 281

QY 225 TMDLAVALGYRKVQVQVIENHILKLFQS 252
Db 282 NSMDLAVALGYRGVQQAIESHLLKLLQN 309

```

RESULT 5
HMT1_HUMAN
ID HMT1_HUMAN STANDARD; PRT; 1267 AA.
AC Q9H9B1; Q8TCN7; Q96F53; Q96JF1; Q96KH4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Histone-lysine N-methyltransferase, H3 lysine-9 specific 5
DE (EC 2.1.1.43) (Histone H3-K9 methyltransferase 5) (H3-K9-HMTase 5)
DE (Euchromatic histone methyltransferase 1) (Eu-HMTase1) (G9a-like
DE protein 1) (GLP1).
GN EUHMTASE1 OR KIAA1876.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1), ENZYMATIC ACTIVITY, AND IDENTIFICATION
RP IN COMPLEX WITH E2F6; TFDP1; MAX; MGA; BAT8; CBX3; RING1; RNF2; MBLR;
RP L3MBTL2 AND YAF2.
RC TISSUE=Cervical carcinoma;
RX MEDLINE=21999559; PubMed=12004135;
RA Ogawa H., Ishiguro K.-I., Gaubatz S., Livingston D.M., Nakatani Y.;
RT "A complex with chromatin modifiers that occupies E2F- and
RT Myc-responsive genes in G0 cells.";
RL Science 296:1132-1136(2002).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Teratocarcinoma;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
RA Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
RA Wakamatsu A., Nakamura Y., Nagahara K., Masuho Y., Oshima A.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 351-1267 FROM N.A. (ISOFORM 1).
RC TISSUE=Brain;
RA Tsuritani K., Ukai Y., Yajima Y., Amemiya C., Yoshimoto M.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 351-1267 FROM N.A. (ISOFORM 3).
RX MEDLINE=22158633; PubMed=12168954;
RA Nakajima D., Okazaki N., Yamakawa H., Kikuno R., Ohara O., Nagase T.;
RT "Construction of expression-ready cDNA clones for KIAA genes: manual
RT curation of 330 KIAA cDNA clones.";
RL DNA Res. 9:99-106(2002).
RN [5]
RP SEQUENCE OF 1150-1267 FROM N.A. (ISOFORM 3), AND TISSUE SPECIFICITY.
RC TISSUE=Brain;
RX MEDLINE=21245130; PubMed=11347906;
RA Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XX.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 8:85-95(2001).
RN [6]
RP SEQUENCE OF 552-1267 FROM N.A. (ISOFORM 1).
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Wang S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ussin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

```

```

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [7]
RP SEQUENCE OF 1056-1267 FROM N.A. (ISOFORM 1).
RC TISSUE=Lymph node;
RA Ansgorge W., Wirkner U., Mewes H.-W., Weil B., Wiemann S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Histone methyltransferase. Methylates Lys-9 of histone
CC H3 (in vitro). H3 Lys-9 methylation represents a specific tag for
CC epigenetic transcriptional repression by recruiting Hpi proteins
CC to methylated histones. Probably targeted to histone H3 by
CC different DNA-binding proteins like E2F6, MGA, MAX and/or DP1.
CC During G0 phase, it probably contributes to silencing of MYC- and
CC E2F-responsive genes, suggesting a role in G0/G1 transition in
CC cell cycle.
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + histone L-lysine =
CC S-adenosyl-L-homocysteine + histone N(6)-methyl-L-lysine.
CC -!- SUBUNIT: Part of the E2F6.com-1 complex in G0 phase composed of
CC E2F6, MGA, MAX, TFDP1, CBX3, BAT8, EUHMTASE1, RING1, RNF2, MBLR,
CC L3MBTL2 and YAF2.
CC -!- SUBCELLULAR LOCATION: Nuclear; associates with euchromatic
CC regions.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Comment=Experimental confirmation may be lacking for some
CC isoforms;
CC Name=1;
CC IsoId=Q9H9B1-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9H9B1-2; Sequence=VSP_002222, VSP_002223;
CC Name=3;
CC IsoId=Q9H9B1-3; Sequence=VSP_002224, VSP_002225;
CC -!- TISSUE SPECIFICITY: Widely expressed.
CC -!- SIMILARITY: Belongs to the histone-lysine methyltransferase
CC family.
CC -!- SIMILARITY: Contains 8 ANK repeats.
CC -!- SIMILARITY: Contains 1 pre-SET domain.
CC -!- SIMILARITY: Contains 1 SET domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AY083210; AAM09024.1; -
CC EMBL; AK022941; BAB14321.1; -
CC EMBL; AB028932; BAB56104.1; -
CC EMBL; AB058779; BAB47505.2; -
CC EMBL; BC011608; AAH11608.2; -
CC EMBL; AL713772; CAD28534.1; ALT_INIT.
CC MIM; 607001; -
CC InterPro; IPR002110; ANK.
CC InterPro; IPR007728; Pre-SET.
CC InterPro; IPR001214; SET.
CC InterPro; IPR003606; Zn2-binding.
CC Pfam; PF00023; ank; 7.
CC Pfam; PF05033; Pre-SET; 1.
CC Pfam; PF00856; SET; 1.
CC PRINTS; PR01415; ANKYRIN.
CC SMART; SM00248; ANK; 7.
CC SMART; SM00468; PreSET; 1.
CC SMART; SM00317; SET; 1.

```

DR PROSITE; PS50088; ANK_REPEAT; 5.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS50867; PRE_SET; 1.
DR PROSITE; PS50280; SET; 1.
KW Transferase; Methyltransferase; Chromatin regulator; Nuclear protein;
KW ANK repeat; Repeat; Alternative splicing.
FT REPEAT 706 735 ANK 1.
FT REPEAT 741 770 ANK 2.
FT REPEAT 774 803 ANK 3.
FT REPEAT 807 837 ANK 4.
FT REPEAT 841 870 ANK 5.
FT REPEAT 874 903 ANK 6.
FT REPEAT 907 936 ANK 7.
FT REPEAT 940 973 ANK 8.
FT DOMAIN 1029 1092 PRE-SET.
FT DOMAIN 1094 1216 SET.
FT DOMAIN 375 378 POLY-GLU.
FT DOMAIN 411 418 POLY-ARG.
FT DOMAIN 1261 1264 POLY-ALA.
FT VARSPLIC 805 825 DAEGSTCLHLAAKKGHYEVVQ -> IQKTSKVYTESQETQR
FT SQTIL (in isoform 2).
FT /FTid=VSP 002222.
FT Missing (in isoform 2).
FT /FTid=VSP 002223.
FT DGEV -> ISSA (in isoform 3).
FT /FTid=VSP 002224.
FT Missing (in isoform 3).
FT /FTid=VSP 002225.
FT G -> E (IN REF. 3 AND 4).
SQ SEQUENCE 1267 AA; 138181 MW; 78FD418FD404F475 CRC64;
Query Match 15.4%; Score 206.5; DB 1; Length 1267;
Best Local Similarity 29.6%; Pred. No. 3.7e-08;
Matches 55; Conservative 33; Mismatches 81; Indels 17; Gaps 3;
QY 62 QAGSSL-----KHSTTLNRQRGNEVSALP-----ATLD-----SLSIHQLAQAQGLD 104
Db 763 QAGANIDTCSEDQRTPLMEAAENHLEAVKYLKAGALVDPKDAEGSTCLHLAAKKGHYE 822
QY 105 QLKEHLRKGDNLVNKPDERGFTPLIWAFAFGEIETVRFLENGADPHILAKERESALSLSA 164
Db 823 VVQYLLSNGQMDVNCQDDGGTTPMIWATEYKHVDLVKLLSKGSDINIRDNEENICLHWA 882
QY 165 STGGYTDIVGLLERRDVIDINIYDNGGTPLLYAVRGNHVKVCVEALLARGADLTTEADSGY 224
Db 883 AFGGCVDAIEILLAAKCDLHAVNIHGDSPHLIAARENRYDCVVLFLSRDSVTLKNKEGE 942
QY 225 TPDMLA 230
Db 943 TPLQCA 948
RESULT 6
ANK2 HUMAN
ID ANK2_HUMAN STANDARD; PRT; 3924 AA.
AC Q01484; Q01485;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ankyrin 2 (Brain ankyrin) (Ankyrin B) (Ankyrin, nonerythroid).
GN ANK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
RC TISSUE=Brain stem;
RX MEDLINE=91302466; PubMed=1830053;
RA Otto E.; Kunimoto M.; McLaughlin T.; Bennett V.;
RT "Isolation and characterization of cDNAs encoding human brain
ankyrins reveal a family of alternatively spliced genes."
RL J. Cell Biol. 114:241-253 (1991).

[2]
REVISIONS.
RA Carpenter S.;
RN Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Brain stem;
RX MEDLINE=94075409; PubMed=8253844;
RA Chan W.; Kordeli E.; Bennett V.;
RT "440-kD ankyrinB: structure of the major developmentally regulated
domain and selective localization in unmyelinated axons."
RL J. Cell Biol. 123:1463-1473 (1993).
[4]
SEQUENCE OF 463-495 FROM N.A.
RX MEDLINE=92009921; PubMed=1833308;
RA Tse W.T.; Menninger J.C.; Yang-Feng T.L.; Francke U.; Sahr K.E.;
RA Lux S.E.; Ward D.C.; Forget B.G.;
RT "Isolation and chromosomal localization of a novel nonerythroid
ankyrin gene."
RL Genomics 10:858-866 (1991).
CC -!- FUNCTION: Attach integral membrane proteins to cytoskeletal
elements. Also bind to cytoskeletal proteins.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=Q01484-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q01484-2; Sequence=VSP_000267, VSP_000268;
CC Name=3;
CC IsoId=Q01484-3; Sequence=VSP_000268;
CC -!- TISSUE SPECIFICITY: Plasma membrane of neurons as well as glial
cells throughout the brain.
CC -!- PTM: Phosphorylated at multiple sites by different protein kinases
and each phosphorylation event regulates the protein's structure
and function (Potential).
CC -!- SIMILARITY: Contains 23 ANK repeats.
CC -!- SIMILARITY: Contains 1 death domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; X56957; CAA40278.1; -
DR EMBL; X56958; CAA40279.2; -
DR EMBL; Z26634; CAB42644.1; -
DR EMBL; M37123; AAA62828.1; -
DR PIR; S37431; S37431.
DR HSSP; P42771; IDC2.
DR Genew; HGNC:493; ANK2.
DR MIM; 106410; -
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000488; Death.
DR InterPro; IPR000906; ZU5.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00791; ZU5; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 22.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00218; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 20.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
KW Cytoskeleton; Alternative splicing; Repeat; ANK repeat;
KW Phosphorylation.
FT REPEAT 63 92 ANK 1.
FT REPEAT 96 125 ANK 2.
FT REPEAT 129 158 ANK 3.
FT REPEAT 162 191 ANK 4.

FT	REPEAT	193	220	ANK 5.
FT	REPEAT	232	261	ANK 6.
FT	REPEAT	265	294	ANK 7.
FT	REPEAT	298	327	ANK 8.
FT	REPEAT	331	360	ANK 9.
FT	REPEAT	364	393	ANK 10.
FT	REPEAT	397	426	ANK 11.
FT	REPEAT	430	459	ANK 12.
FT	REPEAT	463	492	ANK 13.
FT	REPEAT	496	525	ANK 14.
FT	REPEAT	529	558	ANK 15.
FT	REPEAT	562	591	ANK 16.
FT	REPEAT	595	624	ANK 17.
FT	REPEAT	628	657	ANK 18.
FT	REPEAT	661	690	ANK 19.
FT	REPEAT	694	723	ANK 20.
FT	REPEAT	727	756	ANK 21.
FT	REPEAT	760	789	ANK 22.
FT	REPEAT	793	822	ANK 23.
FT	DOMAIN	1773	1950	REPEAT-RICH REGION.
FT	REPEAT	1773	1784	REPEAT A.
FT	REPEAT	1785	1796	REPEAT A.
FT	REPEAT	1797	1808	REPEAT A.
FT	REPEAT	1809	1820	REPEAT A.
FT	REPEAT	1821	1832	REPEAT A.
FT	REPEAT	1833	1844	REPEAT A.
FT	REPEAT	1845	1856	REPEAT A.
FT	REPEAT	1857	1867	REPEAT A.
FT	REPEAT	1868	1879	REPEAT A.
FT	REPEAT	1880	1891	REPEAT A.
FT	REPEAT	1892	1902	REPEAT A.
FT	REPEAT	1903	1914	REPEAT A.
FT	REPEAT	1915	1926	REPEAT A.
FT	REPEAT	1927	1938	REPEAT A.
FT	REPEAT	1939	1950	REPEAT A.
FT	DOMAIN	3536	3620	DEATH.
FT	VARSPLIC	1039	1039	Q -> QFLGKLHLPTAPPPPLNEGESLVSRLQLGPPGTYK (in isoform 2).
FT	VARSPLIC	1444	3528	/FTid=VSP 000267. Missing (in isoform 2 and isoform 3).
FT	CONFLICT	475	476	/FTid=VSP 000268. GQ -> PE (IN REF. 4).
FT	CONFLICT	971	971	I -> S (IN REF. 1).
FT	CONFLICT	3581	3582	QY -> HA (IN REF. 1).
FT	CONFLICT	3586	3586	I -> Y (IN REF. 1).
SQ	SEQUENCE	3924	AA; 430337 MW; 52AC496C428E29D2 CRC64;	
Query Match 15.1%; Score 202.5; DB 1; Length 3924;				
Best Local Similarity 29.3%; Pred. No. 3.4e-07;				
Matches 56; Conservative 43; Mismatches 81; Indels 11; Gaps 4;				
Qy	55	DASVSSPQAGSSLKHSHTLTNRQRGNEVSALPATLDSLSIHQLAAQAGELDLKEHLRKG	114	
Db	5	DAAQKS-DSGEKFNQSSQRRKPKKSDSNA-----SFLRAARAGNLDKVVVEYLKGGI	55	
Qy	115	NLVNKPDERGFTPLIWAFAFGEIETVRFLLWGAADPHILAKERESALSLSLASTGTYTDIVG	174	
Db	56	D-INTCNQGNLHAAKEGHVGLVQELLGRGSSVDSATKKGNTALHIASLAGQAEVVK	114	
Qy	175	LLLRDVIDNIYDWNNGTTPLLYAVRGNHVKCEALLARGADLTTEADSGYTPMDLVALG	234	
Db	115	VLVRGANINAQSQNGFTPLYMAAQENHIDVVKYLLENGANQSTATEDGFTPLAVALQOG	174	
Qy	235	YRK-VQGVEN	244	
Db	175	HNQAVAILLEN	185	
RESULT 7				
ANK3_HUMAN				
ID	ANK3_HUMAN			
AC	Q12955;			
DT	28-FEB-2003 (Rel. 41, Created)			

DT	28-FEB-2003 (Rel. 41, Last sequence update)
DT	15-MAR-2004 (Rel. 43, Last annotation update)
DE	Ankyrin 3 (ANK-3) (Ankyrin G).
GN	ANK3.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Brain stem;
RX	MEDLINE=95138209; PubMed=7836469;
RA	Kordeli B., Lambert S., Bennett V.;
RT	"AnkyrinG. A new ankyrin gene with neural-specific isoforms localized
RT	at the axonal initial segment and node of Ranvier.";
RL	J. Biol. Chem. 270:2352-2359(1995).
CC	!- FUNCTION: Membrane-cytoskeleton linker. The neural-specific
CC	isoforms may participate in the maintenance/targeting of ion
CC	channels and cell adhesion molecules at the nodes of Ranvier and
CC	axonal initial segments.
CC	!- SUBUNIT: Neural-specific isoforms may be a constituent of a
CC	neurofascin/NRCAM/ankyrin G complex.
CC	!- ALTERNATIVE PRODUCTS:
CC	Event=Alternative splicing; Named isoforms=1;
CC	Comment=A number of isoforms are produced;
CC	Name=1; Synonyms=480-kDa isoform;
CC	Isoid=Q12955-1; Sequence=Displayed;
CC	!- TISSUE SPECIFICITY: Expressed in brain and other tissues. Isoform
CC	1 is neural-specific.
CC	!- SIMILARITY: Contains 23 ANK repeats.
CC	!- SIMILARITY: Contains 1 death domain.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; U13616; AAA64834.1; -.
DR	PIR; A55575; A55575.
DR	HSSP; P55273; 1BI8.
DR	Genew; HGNC:494; ANK3.
DR	MIM; 600465; -.
DR	GO; GO:0006605; P:protein targeting; NAS.
DR	InterPro; IPR002110; ANK.
DR	InterPro; IPR000488; Death.
DR	InterPro; IPR000906; ZU5.
DR	Pfam; PF00023; ank; 24.
DR	Pfam; PF00531; death; 1.
DR	Pfam; PF00791; ZU5; 1.
DR	PRINTS; PR01415; ANKYRIN.
DR	SMART; SM00248; ANK; 21.
DR	SMART; SM00005; DEATH; 1.
DR	SMART; SM00218; ZU5; 1.
DR	PROSITE; PS50088; ANK_REPEAT; 21.
DR	PROSITE; PS0297; ANK_REPEAT_REGION; 1.
DR	PROSITE; PS50017; DEATH_DOMAIN; 1.
KW	Cytoskeleton; Alternative splicing; Repeat; ANK repeat.
FT	REPEAT 73 102
FT	REPEAT 106 135
FT	REPEAT 139 168
FT	REPEAT 172 201
FT	REPEAT 203 230
FT	REPEAT 234 263
FT	REPEAT 267 296
FT	REPEAT 300 329
FT	REPEAT 333 362
FT	REPEAT 366 395
FT	REPEAT 399 428
FT	REPEAT 432 461
FT	REPEAT 465 494

FT REPEAT 498 527 ANK 14.
FT REPEAT 531 560 ANK 15.
FT REPEAT 564 593 ANK 16.
FT REPEAT 597 626 ANK 17.
FT REPEAT 630 659 ANK 18.
FT REPEAT 663 692 ANK 19.
FT REPEAT 696 725 ANK 20.
FT REPEAT 729 758 ANK 21.
FT REPEAT 762 791 ANK 22.
FT REPEAT 795 825 ANK 23.
FT DOMAIN 1519 1898 SER-RICH.
FT DOMAIN 4090 4174 DEATH.
SQ SEQUENCE 4377 AA; 480399 MW; F42379E55768B684 CRC64;

Query Match 14.4%; Score 193; DB 1; Length 4377;
Best Local Similarity 28.1%; Pred. No. 2.2e-06;
Matches 50; Conservative 43; Mismatches 75; Indels 10; Gaps 3;

QY 68 KHSTTLNRQNEVSALPATLDSLSIHQLAAQGLDQLKEHLRKGDNLVNKPDERGFTP 127
Db KHRKRSRDRKKSDANA-----SVLRARAGHLEKALDYKNGVD-INICNQGLNA 77

QY 128 LIWASAFGEIETVRFLLEWGADPHILAKERESALSLSASTGGYTDIVGLLLERDVINIYD 187
Db LHLASKEGHEVVSSELLQREANVDATKKGNTALHIASLAGQAEVVKVLVTNGANVNAQS 137

QY 188 WNGGTPLLYAVRGNHVKCVALLARGADLTTEADSGYTPMDLAVAGY-RKVOQVIEN 244
Db QNGFTPLYMAAQENHLEVVKFLLDNGASQSLATEDGFTPLAVALQGHQDVVSLLEN 195

RESULT 8
BAT8 HUMAN
ID BAT8_HUMAN STANDARD; PRT; 1210 AA.
AC Q96KQ7; Q14349; Q96MH5; Q96QD0; Q9UQL8; Q9Y331;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Histone-lysine N-methyltransferase, H3 lysine-9 specific 3
DE (EC 2.1.1.43) (Histone H3-K9 methyltransferase 3) (H3-K9-HMTase 3)
DE (HLA-B associated transcript 8) (G9a) (NG336).
GN BAT8 OR G9A OR NG336.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Rowen L., Qin S., Madan A., Dickhoff R., Dors M., Madan A., Hicks P.,
RA Loretz C., Ratcliffe A., Abbasi N., Shaffer T., Hood L.;
RT "Sequence of the human major histocompatibility complex class III
region.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Hirakawa M., Yamaguchi H., Imai K., Shimada J., Shiina S., Tamiya G.,
RA Oka A., Inoko H.;
RT "Homo sapiens 2,229,817bp genomic DNA of 6p21.3 HLA class I region.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1), ALTERNATIVE SPLICING (ISOFORM 2), AND
RP TISSUE SPECIFICITY.
RX MEDLINE=21564388; PubMed=11707778;
RA Brown S.E., Campbell R.D., Sanderson C.M.;
RT "Novel NG36/G9a gene products encoded within the human and mouse MHC
class III regions.";
RL Mamm. Genome 12:916-924 (2001).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 3).
RC TISSUE=Salivary gland;
RA Oshima A., Takahashi-Fujii A., Tanase T., Imose N., Takeuchi K.,
RA Arita M., Musashino K., Yuuki H., Hara H., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,

RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA Murakawa K., Kanehori K., Sugiyama A., Kawakami B., Suzuki Y.,
RA Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 5-1210 FROM N.A. (ISOFORM 1).
RC TISSUE=Muscle, and Uterus;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Touchman J.W., Green E.D., Dickson M.C.,
RA Blakesley R.W., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywincki M.I., Skalska U., Smailus D.E.,
RA Schnierch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [6]
RP SEQUENCE OF 195-1210 FROM N.A., AND CHARACTERIZATION.
RC TISSUE=Histiocytic lymphoma;
RX MEDLINE=93207535; PubMed=8457211;
RA Milner C.M., Campbell R.D.;
RT "The G9a gene in the human major histocompatibility complex encodes a
novel protein containing ankyrin-like repeats.";
RL Biochem. J. 290:811-818 (1993).
RN [7]
RP ENZYMIC ACTIVITY, AND SUBCELLULAR LOCATION.
RX MEDLINE=21326082; PubMed=11316813;
RA Tachibana M., Sugimoto K., Fukushima T., Shinkai Y.;
RT "Set domain-containing protein, G9a, is a novel lysine-preferring
mammalian histone methyltransferase with hyperactivity and specific
selectivity to lysines 9 and 27 of histone H3.";
RL J. Biol. Chem. 276:25309-25317 (2001).
RN [8]
RP IDENTIFICATION IN COMPLEX WITH E2F6; TFDPI; MAX; MGA; EHM1TASE1; CBX3;
RP RING1; RNF2; MBLR; L3MBTL2 AND YAF2.
RX MEDLINE=21999559; PubMed=12004135;
RA Ogawa H., Ishiguro K.-I., Gaubatz S., Livingston D.M., Nakatani Y.;
RT "A complex with chromatin modifiers that occupies E2F- and
Myc-responsive genes in G0 cells.";
RL Science 296:1132-1136 (2002).
RN [9]
RP FUNCTION: Histone methyltransferase. Preferentially methylates
Lys-9 of histone H3 and Lys-27 of histone H3 (in vitro). H3 Lys-9
methylation represents a specific tag for epigenetic
transcriptional repression by recruiting Hpi proteins to
methylated histones. Probably targeted to histone H3 by different
DNA-binding proteins like E2F6, MGA, MAX and/or Dp1. Also
methylates histone H1 (By similarity).
CC CATALYTIC ACTIVITY: S-adenosyl-L-methionine + histone L-lysine =
S-adenosyl-L-homocysteine + histone N(6)-methyl-L-lysine.
CC SUBUNIT: Part of the E2F6.com-1 complex in G0 phase composed of
E2F6, MGA, MAX, TFDPI, CBX3, BAT8, EHM1TASE1, RING1, RNF2, MBLR,
L3MBTL2 and YAF2.
CC SUBCELLULAR LOCATION: Nuclear; associates with euchromatic
regions. Does not associate with heterochromatin.
CC ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Comment=Additional isoforms seem to exist;
CC Name=1;
CC IsoId=Q96KQ7-1; Sequence=Displayed;

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriques S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP ALTERNATIVE SPLICING (ISOFORM 2).
RX MEDLINE=21564388; PubMed=11707778;
RA Brown S.E., Campbell R.D., Sanderson C.M.;
RT "Novel NG36/G9a gene products encoded within the human and mouse MHC
RT class III regions.";
RL Mamm. Genome 12:916-924(2001).
RN [5]
RP ENZYMATIC ACTIVITY, SUBCELLULAR LOCATION, AND MUTAGENESIS OF ARG-1162.
RX MEDLINE=21326082; PubMed=11316813;
RA Tachibana M., Sugimoto K., Fukushima T., Shinkai Y.;
RT "Set domain-containing protein, G9a, is a novel lysine-preferring
RT mammalian histone methyltransferase with hyperactivity and specific
RT selectivity to lysines 9 and 27 of histone H3.";
RL J. Biol. Chem. 276:25309-25317(2001).
CC -!- FUNCTION: Histone methyltransferase. Preferentially methylates
CC Lys-9 of histone H3 and Lys-27 of histone H3 (in vitro). H3 Lys-9
CC methylation represents a specific tag for epigenetic
CC transcriptional repression by recruiting HP1 proteins to
CC methylated histones. Also methylates histone H1.
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + histone L-lysine =
CC S-adenosyl-L-homocysteine + histone N(6)-methyl-L-lysine.
CC -!- SUBUNIT: Part of the E2F6.com-1 complex in G0 phase composed of
CC E2F6, MGA, MAX, TEP1, CBX3, BAT8, EHM1, RING1, RNF2, MBLR,
CC L3MBTL2 and YAF2 (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear; almost excluded from nucleoli.
CC Associates with euchromatic regions. Does not associate with
CC heterochromatin.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1; Synonyms=G9a-L;
CC IsoId=Q9Z148-1; Sequence=Displayed;
CC Name=2; Synonyms=G9a-S;
CC IsoId=Q9Z148-2; Sequence=VSP_002214, VSP_002215, VSP_002216;
CC -!- MISCELLANEOUS: G9a deficient mice show a higher level of histone
CC H3 with acetylated Lys-9 and/or methylated Lys-4, display severe
CC developmental defects and die within E9.5-E12.5 stages.
CC -!- SIMILARITY: Belongs to the histone-lysine methyltransferase
CC family. Suvar3-9 subfamily.
CC -!- SIMILARITY: Contains 7 ANK repeats.
CC -!- SIMILARITY: Contains 1 pre-SET domain.
CC -!- SIMILARITY: Contains 1 SET domain.
CC -!- SIMILARITY: Contains 1 post-SET domain.
CC -!- CAUTION: NG36 and G9a were originally thought to derive from two
CC separate genes.
CC -!- CAUTION: Ref.1 sequences differ from that shown due to erroneous
CC gene model prediction.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF109906; AAC84164.1; ALT_SEQ.
DR EMBL; AF109906; AAC84165.1; ALT_SEQ.

DR EMBL; AB077209; BAC05482.1; -.
DR EMBL; AB077210; BAC05483.1; -.
DR EMBL; BC025539; AAH25539.1; ALT_INIT.
DR MGD; MGI:2148922; Bat8.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR003616; PostSET.
DR InterPro; IPR007728; Pre-SET.
DR InterPro; IPR001214; SET.
DR InterPro; IPR003606; Zn2-binding.
DR Pfam; PF00023; ank; 6.
DR Pfam; PF05033; Pre-SET; 1.
DR Pfam; PF00856; SET; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 6.
DR SMART; SM00468; PreSET; 1.
DR SMART; SM00317; SET; 1.
DR PROSITE; PS50088; ANK_REPEAT; 5.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS50868; POST_SET; FALSE_NEG.
DR PROSITE; PS50867; PRE_SET; 1.
DR PROSITE; PS50280; SET; 1.
KW Transferase; Methyltransferase; Chromatin regulator; Nuclear protein;
KW ANK repeat; Repeat; Alternative splicing.
FT DOMAIN 352 379 POLY-GLU.
FT REPEAT 702 731 ANK 1.
FT REPEAT 737 766 ANK 2.
FT REPEAT 770 799 ANK 3.
FT REPEAT 803 833 ANK 4.
FT REPEAT 837 866 ANK 5.
FT REPEAT 870 899 ANK 6.
FT REPEAT 903 932 ANK 7.
FT DOMAIN 1025 1088 PRE-SET.
FT DOMAIN 1090 1212 SET.
FT DOMAIN 1217 1233 POST-SET.
FT VARSPPLIC 1 57 Missing (in isoform 2).
FT VARSPPLIC 58 71 /FTId=VSP_002214.
FT VARSPPLIC 426 459 AGLTGPVPVCLPSQ -> MAAAAGAAAAAAAE (in
FT VARSPPLIC 426 459 isoform 2).
FT VARSPPLIC 426 459 /FTId=VSP_002215.
FT MUTAGEN 1162 1162 Missing (in isoform 2).
FT MUTAGEN 1162 1162 /FTId=VSP_002216.
FT MUTAGEN 1165 1168 R-SH: STRONGLY REDUCES HISTONE
FT MUTAGEN 1165 1168 METHYLTRANSFERASE ACTIVITY.
FT MISSING: ABOLISHES HISTONE
FT METHYLTRANSFERASE ACTIVITY AND SUBSEQUENT
FT REPRESSION.
SQ SEQUENCE 1263 AA; 138038 MW; 74DBFF9A36769589 CRC64;
Query Match 14.0%; Score 188; DB 1; Length 1263;
Best Local Similarity 34.5%; Pred. No. 1e-06;
Matches 48; Conservative 19; Mismatches 72; Indels 0; Gaps 0;
QY 91 SLSIHQLAAQAGELDQLKEHLRKGNLKNKPDGERGFTPLIWAFAFGEIETVRFLEWGADP 150
Db 805 STCLHHAAKIGNLEMVSLLLSTGQVDVNAQDSGGWTPIIWAAEHKHIDVIRMLLTRGADV 864
QY 151 HILAKERESALSLASTGGYTDIVGLLLERDVIDINIYDMNGGTPLLIYAVRGNHVKVCVEALL 210
Db 865 TLTDNEENICLHWASFTGSAATAEVLINAQCDLHAVNYHGDTPLHIAARESYHDCVLLFL 924
QY 211 ARGADLTTEADSGYTPMDL 229
Db 925 SRGANFELRNKEGDTAWDL 943
RESULT 10
GASZ MOUSE
ID GASZ MOUSE STANDARD; PRT; 475 AA.
AC Q8VD46; Q9JKQ7;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Germ cell-specific ankyrin, SAM and basic leucine zipper domain

CC Binds to the N-terminus of telomeric TRF1 via the ANK repeats.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic, associated with the Golgi and
CC with juxtanuclear SLC2A4/GLUT4-vesicles. A minor proportion is
CC also found at nuclear pore complexes and around the pericentriolar
CC matrix of mitotic centrosomes. During interphase, a small fraction
CC of TNKS is found in the nucleus, associated with TRF1.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=O95271-1; Sequence=Displayed;
CC Name=2;
CC IsoId=O95271-2; Sequence=VSP_004538, VSP_004539;
CC Note=No experimental confirmation available;
CC -!- TISSUE SPECIFICITY: Ubiquitous; highest levels in testis.
CC -!- PTM: Upon insulin-stimulation, phosphorylated on serine residues
CC by MAPK kinases.
CC -!- PTM: ADP-ribosylated (-auto).
CC -!- SIMILARITY: Belongs to the PARP family.
CC -!- SIMILARITY: Contains 15 ANK repeats.
CC -!- SIMILARITY: Contains 1 sterile alpha motif (SAM) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF082556; AAC79841.1; --
CC EMBL; AF082557; AAC79842.1; --
CC EMBL; AF082558; AAC79843.1; --
CC EMBL; AF082559; AAC79844.1; --
CC HSSP; Q00420; 1AWC.
CC Genew; HGNC:11941; TNKS.
CC MIM; 603303; --
CC GO; GO:0000781; C:chromosome, telomeric region; IDA.
CC GO; GO:0003950; F:NAD ADP-ribosyltransferase activity; IDA.
CC GO; GO:0005515; F:protein binding; IPI.
CC GO; GO:0007004; P:telomerase-dependent telomere maintenance; NAS.
CC InterPro; IPR002110; ANK.
CC InterPro; IPR001660; SAM.
CC Pfam; PF00023; ank; 19.
CC Pfam; PF00536; SAM; 1.
CC PRINTS; PR01415; ANKYRIN.
CC SMART; SM00248; ANK; 17.
CC SMART; SM00454; SAM; 1.
CC PROSITE; PS50088; ANK_REPEAT; 15.
CC PROSITE; PS50297; ANK_REPEAT_REGION; 1.
CC PROSITE; PS50105; SAM_DOMAIN; 1.
CC Transferrase; Glycosyltransferase; NAD; Golgi stack; Telomere;
KW Nuclear protein; Repeat; ANK repeat; ADP-ribosylation;
KW Phosphorylation; Alternative splicing.
FT REPEAT 215 247 ANK 1.
FT REPEAT 248 280 ANK 2.
FT REPEAT 281 313 ANK 3.
FT REPEAT 368 400 ANK 4.
FT REPEAT 401 433 ANK 5.
FT REPEAT 434 466 ANK 6.
FT REPEAT 521 556 ANK 7.
FT REPEAT 557 589 ANK 8.
FT REPEAT 590 622 ANK 9.
FT REPEAT 683 715 ANK 10.
FT REPEAT 716 748 ANK 11.
FT REPEAT 749 781 ANK 12.
FT REPEAT 836 868 ANK 13.
FT REPEAT 869 901 ANK 14.
FT REPEAT 902 934 ANK 15.
FT DOMAIN 1030 1089 SAM.
FT DOMAIN 1176 1327 PARP.
FT DOMAIN 9 14 POLY-HIS.
FT DOMAIN 27 34 POLY-PRO.
FT DOMAIN 128 134 POLY-SER.

FT DOMAIN 137 145 POLY-SER.
FT VARSPLIC 641 643 EST -> GHS (in isoform 2).
FT VARSPLIC 644 1327 /FTId=VSP_004538.
FT MUTAGEN 1184 1184 Missing (in isoform 2).
FT MUTAGEN 1291 1291 /FTId=VSP_004539.
FT MUTAGEN 1291 1291 H->A: LOSS OF ACTIVITY; WHEN ASSOCIATED
FT MUTAGEN 1291 1291 WITH A-1291.
FT MUTAGEN 1291 1291 E->A: LOSS OF ACTIVITY; WHEN ASSOCIATED
FT MUTAGEN 1291 1291 WITH A-1184.
SQ SEQUENCE 1327 AA; 142010 MW; E14DE985C710B957 CRC64;

Query Match 13.9%; Score 187; DB 1; Length 1327;
Best Local Similarity 24.3%; Pred. No. 1.3e-06;
Matches 72; Conservative 36; Mismatches 112; Indels 76; Gaps 9;

QY 18 ASELGDPEDPGEEAAGSDTV-VLSLFPCTPEPVNPEPDASVSPQAGSSLKHSHTLTNR 76
Db 137 SSSSSSPSPGSSLAESPEAAGVSSTAPLPGGAAGP-----GTGVFAVSGALRE---LLEA 189
QY 77 QRGNEVSALPATLDSLSI-----HQLAAQAGELDQKHEHLRKGDNLVNKPDER 123
Db 190 CRNGDVSRVKRLVDAANVNNAKDMAGRKSSPLHFAAGFGRKDVVEHLQMGAN-VHARDG 248
QY 124 GFTPLIWASAFGEIETVRFL-----EW-----GADP 150
Db 249 GLIPLHNACSFHAEVVSLLCCQADPNARDNNYTPHAAIKGIDVICIVLLQHGADP 308
QY 151 HILAKERESALSAS-----TGGY-----TDIVGLLLERDVIDINIYDWN 189
Db 309 NIRNTDCKSALDLADPSAKAVLTGEYKKDELLEAARSNGEEKLMALLTPLNVNCHASDGR 368
QY 190 GGTPLLYAVRGNHVKCVEALLARGADLTTEADSGYTPMDLAVALGYRKVQVVIENH 245
Db 369 KSTPLHLAAGYNRVRIVQLLQHGADVHAKDKGLVPLHNACSYGHYEVTETLLKH 424

RESULT 12
ANK1 HUMAN
ID ANK1 HUMAN STANDARD; PRT; 1880 AA.
AC P16157;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Ankyrin 1 (Erythrocyte ankyrin) (Ankyrin R).
GN ANK1 OR ANK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND VARIANTS.
RC TISSUE=Hematopoietic;
RX MEDLINE=90158830; PubMed=2137557;
RA Lux S.E., John K.M., Bennett V.;
RT "Analysis of cDNA for human erythrocyte ankyrin indicates a repeated
RT structure with homology to tissue-differentiation and cell-cycle
RT control proteins."
RL Nature 344:36-42(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90175370; PubMed=1689849;
RA Lambert S., Yu H., Prchal J.T., Lawler J., Ruff P., Speicher D.,
RA Cheung M.C., Kan Y.W., Palek J.;
RT "cDNA sequence for human erythrocyte ankyrin."
RL Proc. Natl. Acad. Sci. U.S.A. 87:1730-1734(1990).
RN [3]
RP VARIANT HS ILE-462.
RX MEDLINE=96225450; PubMed=8640229;
RA Eber S.W., Gonzalez J.M., Lux M.L., Scarpa A.L., Tse W.T.,
RA Dornwell M., Herbers J., Kugler W., Oezcan R., Pekrun A.,
RA Gallagher P.G., Schroeter W., Forget B.G., Lux S.E.;
RT "Ankyrin-1 mutations are a major cause of dominant and recessive
RT hereditary spherocytosis.";

RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -!- FUNCTION: Acts as a regulatory subunit of the 26S proteasome which
CC is involved in the ATP-dependent degradation of ubiquitinated
CC proteins.
CC -!- SUBUNIT: Component of the PA700 complex.
CC -!- SIMILARITY: Contains 5 ANK repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB022022; BAA36969.1; -;
CC EMBL; AK009068; BAB26053.1; -;
CC HSSP; P42773; I1HB.
CC MGD; MGI:1858898; Psm10.
CC InterPro; IPR002110; ANK.
CC Pfam; PF00023; ank; 5.
CC PRINTS; PR01415; ANKYRIN.
CC SMART; SM00248; ANK; 5.
CC PROSITE; PS50088; ANK_REPEAT; 5.
CC PROSITE; PS50297; ANK_REPEAT; 1.
CC Proteasome; ANK repeat; Repeat.
FT REPEAT 39 68 ANK 1.
FT REPEAT 72 101 ANK 2.
FT REPEAT 105 134 ANK 3.
FT REPEAT 138 167 ANK 4.
FT REPEAT 171 200 ANK 5.
FT CONFLICT 101 101 A -> S (IN REF. 1).
FT CONFLICT 122 122 A -> S (IN REF. 1).
FT CONFLICT 226 226 G -> S (IN REF. 1).
SQ SEQUENCE 231 AA; 25053 MW; BCE7B9A79C8CF58B CRC64;

Query Match 13.6%; Score 182.5; DB 1; Length 231;
Best Local Similarity 31.4%; Pred. No. 2.9e-07;
Matches 54; Conservative 27; Mismatches 86; Indels 5; Gaps 1;
QY 89 LDSLSIHQLAAGQELDQLKEHLRKGNDLVNKPDERGFTPLIWASAFGEIETVRFLLEWGA 148
Db 5 VSNIMICNLAYSGLDELKERILADKSLATRDQDSRTALHWACSAGHTEIVEFLQLGV 64
QY 149 DPHILAKERESALSASTGGYTDIVGLLLERDVIDINIYDNGTPLLVAVRGNHVKVCEA 208
Db 65 PVNDKDDAGWSPLHIAASAGRDEIVKALLVKGAVNAVQNGCTPLHYAASKNRHEIAVM 124
QY 209 LLARGADLTTEADSGYTPMDLAVALGYRKVQVQVIENHILKLFQSNLVPADPE 260
Db 125 LLEGANPDADKHDYDATAMHRAAKGNLKMV-----HILFYKASTNIQDTE 171

Query Match 13.6%; Score 182.5; DB 1; Length 231;
Best Local Similarity 31.4%; Pred. No. 2.9e-07;
Matches 54; Conservative 27; Mismatches 86; Indels 5; Gaps 1;
QY 89 LDSLSIHQLAAGQELDQLKEHLRKGNDLVNKPDERGFTPLIWASAFGEIETVRFLLEWGA 148
Db 5 VSNIMICNLAYSGLDELKERILADKSLATRDQDSRTALHWACSAGHTEIVEFLQLGV 64
QY 149 DPHILAKERESALSASTGGYTDIVGLLLERDVIDINIYDNGTPLLVAVRGNHVKVCEA 208
Db 65 PVNDKDDAGWSPLHIAASAGRDEIVKALLVKGAVNAVQNGCTPLHYAASKNRHEIAVM 124
QY 209 LLARGADLTTEADSGYTPMDLAVALGYRKVQVQVIENHILKLFQSNLVPADPE 260
Db 125 LLEGANPDADKHDYDATAMHRAAKGNLKMV-----HILFYKASTNIQDTE 171

RESULT 15
YB23_HUMAN
ID YB23_HUMAN STANDARD; PRT; 768 AA.
AC Q9ULJ7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein KIAA1223 (Fragment).
GN KIAA1223.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20039619; PubMed=10574462;
RA Nagase T., Ishikawa K.-I., Kikuno R., Hirose M., Nomura N.,
RA Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XV.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 6:337-345(1999).
CC -!- SIMILARITY: Contains at least 14 ANK repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB033049; BAA86537.1; -;
CC HSSP; P42771; IDC2.
CC InterPro; IPR002110; ANK.
CC Pfam; PF00023; ank; 14.
CC PRINTS; PR01415; ANKYRIN.
CC SMART; SM00248; ANK; 13.
CC PROSITE; PS50088; ANK_REPEAT; 13.
CC PROSITE; PS50297; ANK_REPEAT; 1.
CC Hypothetical protein; Repeat; ANK repeat.
FT NON TER 1 1
FT REPEAT 1 11 ANK 1.
FT REPEAT 15 44 ANK 2.
FT REPEAT 48 82 ANK 3.
FT REPEAT 86 115 ANK 4.
FT REPEAT 119 148 ANK 5.
FT REPEAT 152 181 ANK 6.
FT REPEAT 185 214 ANK 7.
FT REPEAT 218 247 ANK 8.
FT REPEAT 251 280 ANK 9.
FT REPEAT 284 313 ANK 10.
FT REPEAT 317 346 ANK 11.
FT REPEAT 350 379 ANK 12.
FT REPEAT 383 412 ANK 13.
FT REPEAT 416 446 ANK 14.
SQ SEQUENCE 768 AA; 82819 MW; 2913B69BE2DFE06D CRC64;

Query Match 13.5%; Score 181.5; DB 1; Length 768;
Best Local Similarity 36.2%; Pred. No. 1.7e-06;
Matches 46; Conservative 25; Mismatches 51; Indels 5; Gaps 1;
QY 117 VNKPDDEGFTPLIWASAFGEIETVRFLLEWGADPHILAKERESALSALA-----STGGYTD 171
Db 9 VNKADNEGRTALIAAAYMGHREIVEHLLDHGAENVHEDVDGRTALSVAALCVPASKGHAS 68
QY 172 IVGLLLERDVIDINIYDNGTPLLVAVRGNHVKVCEALLARGADLTTEADSGYTPMDLAV 231
Db 69 VVSLLDIRGAENVHCDKDGMTPLLVAAAYEGHVVDVVDLLEGGADVDTDNGRTPLLA 128
QY 232 ALGYRKV 238
Db 129 SMGHASV 135

Query Match 13.5%; Score 181.5; DB 1; Length 768;
Best Local Similarity 36.2%; Pred. No. 1.7e-06;
Matches 46; Conservative 25; Mismatches 51; Indels 5; Gaps 1;
QY 117 VNKPDDEGFTPLIWASAFGEIETVRFLLEWGADPHILAKERESALSALA-----STGGYTD 171
Db 9 VNKADNEGRTALIAAAYMGHREIVEHLLDHGAENVHEDVDGRTALSVAALCVPASKGHAS 68
QY 172 IVGLLLERDVIDINIYDNGTPLLVAVRGNHVKVCEALLARGADLTTEADSGYTPMDLAV 231
Db 69 VVSLLDIRGAENVHCDKDGMTPLLVAAAYEGHVVDVVDLLEGGADVDTDNGRTPLLA 128
QY 232 ALGYRKV 238
Db 129 SMGHASV 135

Search completed: April 29, 2004, 06:24:35
Job time : 19 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 29, 2004, 06:16:23 ; Search time 45 Seconds
(without alignments)
1822.994 Million cell updates/sec

Title: US-09-840-243C-11
Perfect score: 1341
Sequence: 1 MELTQPAEDLIQTQTPASE.....VIENHILKLFQSNLVPADPE 260

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriaph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	611.5	45.6	296	13 Q7T331	Q7t331 brachydanio
2	439	32.7	144	11 Q8C4M9	Q8c4m9 mus musculu
3	274	20.4	234	5 Q9VL58	Q9vl58 drosophila
4	219.5	16.4	426	16 Q8YTG9	Q8ytg9 anabaena sp
5	215	16.0	1219	11 Q8C8R3	Q8c8r3 mus musculu
6	214	16.0	737	4 Q8NDR5	Q8ndr5 homo sapien
7	214	16.0	1786	5 Q17344	Q17344 caenorhabdi
8	214	16.0	1809	5 Q17487	Q17487 caenorhabdi
9	214	16.0	1815	5 Q17488	Q17488 caenorhabdi
10	214	16.0	1841	5 Q8MQG0	Q8mqg0 caenorhabdi
11	214	16.0	1867	5 Q17486	Q17486 caenorhabdi
12	214	16.0	2039	5 Q17489	Q17489 caenorhabdi
13	214	16.0	6994	5 Q17343	Q17343 caenorhabdi
14	214	16.0	6994	5 Q17490	Q17490 caenorhabdi
15	213.5	15.9	917	11 Q8BHB2	Q8bbh2 mus musculu
16	212	15.8	1762	11 Q9EQG6	Q9eqg6 rattus norv

17	212	15.8	1777	4 Q9ULH0	Q9ulh0 homo sapien
18	210.5	15.7	1159	5 Q9NCP8	Q9ncp8 drosophila
19	210.5	15.7	1599	11 Q99NH0	Q99nh0 mus musculu
20	209.5	15.6	2486	4 O75179	O75179 homo sapien
21	208	15.5	1693	11 Q80TG7	Q80tg7 mus musculu
22	207.5	15.5	1031	4 Q9UF42	Q9uf42 homo sapien
23	207.5	15.5	1715	11 Q9ERD4	Q9erd4 rattus norv
24	205.5	15.3	321	5 Q9VQ11	Q9vqi1 drosophila
25	205	15.3	448	11 Q8BTQ7	Q8btq7 mus musculu
26	205	15.3	469	11 Q80XC9	Q80xc9 mus musculu
27	205	15.3	505	11 Q8BXZ6	Q8bxz6 mus musculu
28	205	15.3	655	11 Q9CZK6	Q9czk6 mus musculu
29	205	15.3	655	11 Q80X46	Q80x46 mus musculu
30	202.5	15.1	765	4 Q8NFD2	Q8nfd2 homo sapien
31	201.5	15.0	1680	13 Q7T163	Q7t163 brachydanio
32	201	15.0	616	4 Q96GK0	Q96gk0 homo sapien
33	199	14.8	454	11 Q8C445	Q8c445 mus musculu
34	199	14.8	1604	5 Q9W2I0	Q9w2i0 drosophila
35	198.5	14.8	1089	5 Q7YU92	Q7yu92 drosophila
36	198.5	14.8	1863	4 Q7Z3L5	Q7z3l5 homo sapien
37	198	14.8	418	11 Q8BNC1	Q8bnc1 mus musculu
38	198	14.8	627	4 Q8WY90	Q8wy90 homo sapien
39	198	14.8	1486	4 Q8TEF1	Q8tef1 homo sapien
40	198	14.8	2542	4 Q8IWZ3	Q8i wz3 homo sapien
41	198	14.8	2617	4 Q8IWZ2	Q8i wz2 homo sapien
42	198	14.8	4001	5 Q8WRQ7	Q8wrq7 drosophila
43	198	14.8	4001	5 Q9VCA8	Q9vca8 drosophila
44	197	14.7	686	4 Q9NXF0	Q9nxf0 homo sapien
45	197	14.7	843	11 P97582	P97582 rattus norv

ALIGNMENTS

RESULT 1

Q7T331	Q7T331	PRELIMINARY;	PRT;	296 AA.
AC	Q7T331;			
DT	01-OCT-2003 (TREMBLrel. 25, Created)			
DT	01-OCT-2003 (TREMBLrel. 25, Last sequence update)			
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)			
DE	Hypothetical protein.			
OS	Brachydanio rerio (Zebrafish) (Danio rerio).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;			
OC	Cyprinidae; Danio.			
OX	NCBI_TaxID=7955;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Kidney;			
RX	MEDLINE=22388257; PubMed=12477932;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,			
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,			
RA	Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,			
RA	Jones S.J., Marra M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length human			
RT	and mouse cDNA sequences."			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).			
RN	[2]			
RP	SEQUENCE FROM N.A.			

RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC053275; AAH53275.1; --
KW Hypothetical protein.
SQ SEQUENCE 296 AA; 32255 MW; 2B0EE1F547044710 CRC64;

Query Match 45.6%; Score 611.5; DB 13; Length 296;
Best Local Similarity 46.1%; Pred. No. 8.1e-44;
Matches 135; Conservative 44; Mismatches 71; Indels 43; Gaps 6;

QY 1 MELTQPAEDLIQTQTPASELGDPEDPGEAAAGSDTVVLSLFP---CT----- 46
Db 1 MEGLCPLPEVTGKVPESVSIGSTEDASSQAMGIKFIPLNRFDMNVCSRFSVKSLSNEEDS 60

QY 47 ---PEPVNPEPD-ASV-----SSPQAG-----SSLKHSSTTLTNRQR 78
Db 61 KNIQDQVNSDLEVASVLFKAECNIQTSPSPGIQVRHVYTPSTTKHFSPKQSTTLTNKHR 120

QY 79 GNEVSALPATLDSLSIHQLAAQGLDQLKEHLRKGDNLVNKPDERGFTPLIWASAFGEIE 138
Db 121 GNEVSTPLLVHLSIHQLAAQGMVFLASRIEQ-ESVINLQDEEGFTPLMWAAAHGQIA 179

QY 139 TVRFLEWGADPHILAKERESALSASTGGTYTDIVGLLLERDVIDINIYDNGGTPLLIYAV 198
Db 180 VVEFLQSGADPHLLAKGRESALSACSKGYTDIVRMLIDCGVDVNEYDNGGAPLLIYAV 239

QY 199 RGNHVKVCVEALLARGADLTTEADSGYTPMDLAVLGYRKVQVQVNIHILKLFQ 251
Db 240 HGNHVRVCVEILLESQSDPTMESDSGFNAMDVAVAMGHRNVQVLEAHULKLLQ 292

RESULT 2
Q8C4M9 PRELIMINARY; PRT; 144 AA.
ID Q8C4M9
AC Q8C4M9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Ankyrin repeat.
GN ANKRA2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs";
RL Nature 420:563-573(2002).
DR EMBL; AK081695; BAC38296.1; --
DR MGD; MGI:1915808; Ankra2.
DR GO; GO:0005624; C:membrane fraction; IDA.
DR GO; GO:0005515; F:protein binding; IPI.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ank; 3.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 3.
DR PROSITE; PS50088; ANK REPEAT; 3.
DR PROSITE; PS50297; ANK REP REGION; 1.
SQ SEQUENCE 144 AA; 15733 MW; B6ACEB1554187899 CRC64;

Query Match 32.7%; Score 439; DB 11; Length 144;
Best Local Similarity 68.3%; Pred. No. 1.4e-29;
Matches 84; Conservative 13; Mismatches 26; Indels 0; Gaps 0;

QY 114 DNLVNKPDERGFTPLIWASAFGEIETVRFLEWGADPHILAKERESALSASTGGTYTDIV 173
Db 11 ENVINHTEDEGTPLMWAAAHGQIAVVEFLQNGADPQLLGKRESALSACSKGYTDIV 70

QY 174 GLLLRDVIDINIYDNGGTPLLYAVRGNHVKVCVEALLARGADLTTEADSGYTPMDLAVL 233
Db 71 KMLLDGVDVNEYDNGGTPLLYAVHGNHVKVCVKMLLENGADPTIETDTSYNSMDLAVL 130

QY 234 GYR 236
Db 131 GYR 133

RESULT 3
Q9VL58 PRELIMINARY; PRT; 234 AA.
ID Q9VL58
AC Q9VL58;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CG5846 protein (LP07441p).
GN CG5846.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny L.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,

RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003626; AAF52838.2; -
DR EMBL; AY061547; AAL29095.1; -
DR HSSP; P80144; 2MYO.
DR FlyBase; FBgn032171; CG5846.
DR InterPro; IPR02110; ANK.
DR Pfam; PF00023; ank; 4.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 3.
DR PROSITE; PS50088; ANK_REPEAT; 2.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
KW ANK repeat; Repeat.
SQ SEQUENCE 234 AA; 24838 MW; A4E8DFBF97BF0514 CRC64;

Query Match 20.4%; Score 274; DB 5; Length 234;
Best Local Similarity 33.8%; Pred. No. 3.3e-15;
Matches 80; Conservative 37; Mismatches 96; Indels 24; Gaps 6;

QY 13 TQQTASELGDPEDEEADGSDTWVLSLFPCTPEPVNPEPDASVSSPQAGSSLKHSST 72
Db 7 TIQTNAN---SDDDEGVSAPTSMLVLDKRSKSAFLPYRPQ-----STV 47
QY 73 LTNRQRGN-EVSALPATLDSLSIHQLAAQGL--DQLKEHLRKGDNLVNKPDERGFTPLI 129
Db 48 LTNLQRGNTTEATFCPVEV-SLSFHERAGQGEITEEQVAERARQONIDYK-DAHGFALH 105
QY 130 WASAFGEIETVRFLLEWGADPHILAKERESALSASTGGYTDIVGLLLERDVVDINIYDWN 189
Db 106 WAASYGQLVSVQLLVAAAGANVTMAPDLISPLLLAAAGGHNEIVRFLLEHGADSGHMDIV 165
QY 190 GGTPLLYAVRGNHVKVEALLARGADLTTEADSGYTPMDLAVALGYRKVQVVIENHI 246
Db 166 GNTALMYAAGNHPTCTNELLAKDLDSATNEDGDTAYSLAVEHGAHLAQALLEQYM 222

RESULT 4
Q8YTG9 PRELIMINARY; PRT; 426 AA.
AC Q8YTG9;
DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
DE Hypothetical protein All2748.
GN ALL2748.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpō S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
cyanobacterium Anabaena sp. strain PCC 7120."
RL DNA Res. 8:205-213(2001).
DR EMBL; AP003590; BAB74447.1; -
DR PIR; AE2149; AE2149.
DR InterPro; IPR02110; ANK.
DR Pfam; PF00023; ank; 12.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 13.
DR PROSITE; PS50088; ANK_REPEAT; 11.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 426 AA; 43800 MW; CF0B9E5D43ACAA03 CRC64;

Query Match 16.4%; Score 219.5; DB 16; Length 426;
Best Local Similarity 30.7%; Pred. No. 3.5e-10;
Matches 62; Conservative 34; Mismatches 71; Indels 35; Gaps 4;

QY 62 QAGSSLKHSSTL-----TNRQRGNEVSALPATLDSLSIHQLAAQGLDQLK 107
Db 234 QDGESALHLATVEGYVDVVQVLLNQGANQIKNKLGDTPLLVAALQGH-----DQIV 285
QY 108 EHLRK-----GDNLVNKPDERGFTPLI WASAFGEIETVRFLLEWGADPHILAKERESAL 161
Db 286 ETLKYGANVHGDNL-----GETPLTLAASQCHTATVRILLDYGANANIPASDGKTAL 338
QY 162 SLASTGGYTDIVGLLLERDVVDINIYDWNCGTPLLVAVRGNHVKVEALLARGADLTTEAD 221
Db 339 IKATERNHPGVIQLLAKGANVYQDSVGATALIWAASGGYNKVQVILLEGADTNLKNR 398
QY 222 SGYTPMDLAVALGYRKVQVQVIE 243
Db 399 GGYTALMIAEFNGFRSIVQILK 420

RESULT 5
Q8C8R3 PRELIMINARY; PRT; 1219 AA.
AC Q8C8R3;
DT 01-MAR-2003 (TremBLrel. 23, Created)
DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Ankyrin 2 (Fragment).
GN ANK2 OR AI835472.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Retina;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK044634; BAC32012.1; -
DR MGD; MGI:88025; Ank2.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000906; ZUS.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00791; ZUS; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 23.
DR SMART; SM00218; ZUS; 1.
DR PROSITE; PS50088; ANK_REPEAT; 20.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
FT NON_TER 1219 1219
SQ SEQUENCE 1219 AA; 131201 MW; AFBD462967AD6184 CRC64;

Query Match 16.0%; Score 215; DB 11; Length 1219;
Best Local Similarity 27.9%; Pred. No. 4.1e-09;
Matches 55; Conservative 43; Mismatches 79; Indels 20; Gaps 3;
QY 49 PVNPEPDASVSSPQAGSSLKHSSTLTNQRGNEVSALPATLDSLSIHQLAAQGLDQLKE 108
Db 4 PTSPGEGGACTQNPPIRQSDS-----NASFLRAARAGNLDKVE 45
QY 109 HLRKGDNLVNKPDERGFTPLI WASAFGEIETVRFLLEWGADPHILAKERESALSASTGG 168
Db 46 YLKGKGD-INTCNQGNLHLLAAKEGHVGLVQELLGRGSSVDSATKKGNTALHIALAG 104
QY 169 YTDIVGLLLERDVVDINIYDWNCGTPLLVAVRGNHVKVEALLARGADLTTEADSGYTPMD 228
Db 105 QAEVVKVLVKEGNINAQSQNGFTPLVMAAQENHIDVVKYLLENGANQSTATEDGFTPLA 164
QY 229 LAVALGYRK-VQQVIEN 244
Db 165 VALQQGHNOAVAILLEN 181

SQ SEQUENCE 1815 AA; 198956 MW; 3ECC6E310915C915 CRC64;
Query Match 16.0%; Score 214; DB 5; Length 1815;
Best Local Similarity 27.3%; Pred. No. 8.9e-09;
Matches 62; Conservative 36; Mismatches 87; Indels 42; Gaps 4;
QY 19 SELGDPEDPGEAAAGSDTVVLSLFPCTPEPVNPEPDASVSSPQAGSSLKHSHTTLNRQR 78
Db 2 SNEGDPQPQQQPPESQEVQA-----PAAPEPGRAE----- 32
QY 79 GNEVSALPATLDSLSIHQAAQAGELDQLKEHLRKGDNLVNKPDERGFTPLIWASAFGEIE 138
Db 33 -----GSASFLRAARAGDLEKVLLELRAGTD-INTSNANGLNSHLASKEGHSE 80
QY 139 TVRFLLEWGADPHILAKERESALSASTGGYTDIVGLLLERDVIDINIDWNGGTPLLYAV 198
Db 81 VVRELKQQAQVDAATRKGTALHIALAGQSLIVTILVENGANVNVQSVNGFTPLYMAA 140
QY 199 RGNHVKVEALLARGADLTTEADSGYTPMDLAVAGY-RKVQQVNIEN 244
Db 141 QENHEEVVKYLLKHGANQALSTEDGFTPLAVALQQGHDRVAVLLEN 187

RESULT 10

Q8MQG0 ID Q8MQG0 PRELIMINARY; PRT; 1841 AA.
AC Q8MQG0;
DT 01-OCT-2002 (TremBLrel. 22, Created)
DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
DE UNC-44 protein (corresponding sequence B0350.2e).
GN B0350.2 OR UNC-44.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA Waterston R.;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Gattung S.;
RT "The sequence of C. elegans cosmid B0350.";
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; U50071; AAM75382.1; -.
DR WormPep; B0350.2e; CE31147.
DR GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000488; Death.
DR InterPro; IPR001360; Glyco_hydro_1.
DR InterPro; IPR000906; ZU5.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00791; ZU5; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 23.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00218; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 22.
DR PROSITE; PS50297; ANK_REPEAT; 1.

DR PROSITE; PS50017; DEATH_DOMAIN; 1.
DR PROSITE; PS00572; GLYCOSYL_HYDROL_F1_1; 2.
KW ANK repeat; Repeat.
SQ SEQUENCE 1841 AA; 201234 MW; 24CC6020DDCE0FC8 CRC64;
Query Match 16.0%; Score 214; DB 5; Length 1841;
Best Local Similarity 27.3%; Pred. No. 9.1e-09;
Matches 62; Conservative 36; Mismatches 87; Indels 42; Gaps 4;
QY 19 SELGDPEDPGEAAAGSDTVVLSLFPCTPEPVNPEPDASVSSPQAGSSLKHSHTTLNRQR 78
Db 2 SNEGDPQPQQQPPESQEVQA-----PAAPEPGRAE----- 32
QY 79 GNEVSALPATLDSLSIHQAAQAGELDQLKEHLRKGDNLVNKPDERGFTPLIWASAFGEIE 138
Db 33 -----GSASFLRAARAGDLEKVLLELRAGTD-INTSNANGLNSHLASKEGHSE 80
QY 139 TVRFLLEWGADPHILAKERESALSASTGGYTDIVGLLLERDVIDINIDWNGGTPLLYAV 198
Db 81 VVRELKQQAQVDAATRKGTALHIALAGQSLIVTILVENGANVNVQSVNGFTPLYMAA 140
QY 199 RGNHVKVEALLARGADLTTEADSGYTPMDLAVAGY-RKVQQVNIEN 244
Db 141 QENHEEVVKYLLKHGANQALSTEDGFTPLAVALQQGHDRVAVLLEN 187

RESULT 11

Q17486 ID Q17486 PRELIMINARY; PRT; 1867 AA.
AC Q17486; O02516;
DT 01-NOV-1996 (TremBLrel. 01, Created)
DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
DE C. ELEGANS ankyrin-related UNC-44 (GB:U21734).
GN B0350.2 OR UNC-44.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A.,
Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Fulton L.,
Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,
Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans.";
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Gattung S.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Waterston R.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=N2;
RX MEDLINE=95263663; PubMed=7744957;
RA Otsuka A.J., Franco R., Yang B., Shim K., Tang L.Z., Zhang Y.,
Boontrakulpoontawe P., Jeyaprakash A., Hedgecock E., Wheaton V.I.,
Sobery A.;
RT "An ankyrin-related gene (unc-44) is necessary for proper axonal
guidance in Caenorhabditis elegans.";
RT J. Cell Biol. 129:1081-1092(1995).

Search completed: April 29, 2004, 06:26:46
Job time : 48 secs